Evaluating onco-geriatric scores and medication risks to improve cancer care for older patients

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

IMKE ORTLAND

aus

Quakenbrück

Bonn 2019

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn elektronisch publiziert.

https://nbn-resolving.org/urn:nbn:de:hbz:5-58042

Erstgutachter: Prof. Dr. Ulrich Jaehde

Zweitgutachter: Prof. Dr. Andreas Jacobs

Tag der Promotion: 28. Februar 2020

Erscheinungsjahr: 2020

Danksagung

Auf dem Weg zur Promotion haben mich viele Menschen begleitet und in ganz unterschiedlicher Weise unterstützt. All diesen Menschen möchte ich an dieser Stelle ganz herzlich danken.

Mein aufrichtiger Dank gilt meinem Doktorvater Prof. Dr. Ulrich Jaehde für das in mich gesetzte Vertrauen, sowie für die Überlassung dieses spannenden Dissertationsthemas. Die uneingeschränkte Unterstützung, wertvollen Diskussionen und die mitreißende Begeisterung für die Wissenschaft haben mich während aller Phasen der Dissertation stets motiviert, unterstützt und sehr viel Wertvolles gelehrt.

Prof. Dr. Andreas Jacobs danke ich herzlich für die Initiierung dieses interessanten Projekts, für die stetige Begeisterung und Unterstützung, sowie für das mir entgegengebrachte Vertrauen. Die ausgezeichnete Zusammenarbeit mit dem Johanniter Krankenhaus Bonn hat ganz maßgeblich zum Gelingen dieser Arbeit beigetragen. Auch danke ich Prof. Dr. Andreas Jacobs herzlich für die Bereitschaft, das Koreferat dieser Arbeit zu übernehmen.

Ebenfalls danke ich herzlich Prof. Dr. Yon-Dschun Ko für seine fortwährende Motivation und seinen Einsatz, sowie für das mir geschenkte Vertrauen, dieses Projekt am Johanniter Krankenhaus zu realisieren. Ebenfalls danke ich Prof. Dr. Yon-Dschun Ko herzlich für die Beteiligung an der Prüfungskommission.

Prof. Dr. Gerd Bendas danke ich ganz herzlich für die Beteiligung an der Prüfungskommission.

Für das Korrekturlesen gilt mein großer Dank Corinna Vučur und Anna Barnert – für die Zeit, die sie für mich geopfert haben und insbesondere für die wertvollen Kommentare und konstruktive Kritik.

Unserer Masterstudentin Monique Mendel Ott danke ich von Herzen für die engagierte Vorarbeit für das Projekt. Dies hat mir den Einstieg in das Thema Onkogeriatrie und in die Arbeit am Johanniter Krankenhaus sehr erleichtert.

Den Mitarbeitern des Johanniter Krankenhauses Bonn danke ich ganz besonders für die freundliche Aufnahme auf den Stationen und in der onkologischen Ambulanz, ihr Interesse an dem Projekt und die engagierte Unterstützung trotz des oftmals intensiven Klinikalltags. Auch

gebührt mein großer Dank den onkologischen Praxen des Bonns/Rhein-Sieg-Kreises für die freundliche Unterstützung bei der Datenerhebung.

Insbesondere danke ich von ganzem Herzen allen Patienten für ihr entgegengebrachtes Vertrauen und die Bereitwilligkeit, sich trotz ihrer schwierigen Situation die Zeit zu nehmen, an meinem Projekt teilzunehmen.

Ich bedanke mich zudem herzlich bei der Herbert-Worch-Stiftung für die finanzielle Förderung dieses Projekts.

Ganz herzlich danke ich meinen Freunden und Kollegen des AK Jaehde für die wundervollen, gemeinsamen Jahre. Durch die warme Atmosphäre, die inspirierenden fachlichen Diskussionen, die stets freundschaftliche Zusammenarbeit, gegenseitige Hilfe und die gute Laune im Team, bin ich jeden Tag mit Freude zum Institut gekommen. Ganz besonders danke ich meinen Weggefährten im „Bungalow“ und in unserem „Lieblingsbüro“ für die großartige Zeit – Patricia, Kerstin, Achim und Eduard.

Schlussendlich danke ich aus tiefstem Herzen meiner Familie und meinen Freunden. Insbesondere meinen Eltern, für ihre bedingungslose Unterstützung an jedem Tag meines Lebens sowie ihre stets aufmunternden Worte und meinen Schwestern, die mich immer gelehrt haben, das Leben mit Humor zu nehmen. Besonders danke ich Sören für seine Anteilnahme, sein Verständnis und seinen Glauben an mich während des langen Wegs bis zur Vollendung der Promotion.

Für meine Eltern

Table of contents

Abbreviations

ABDA	Federal Union of German Associations of Pharmacists (Bundesvereinigung Deutscher Apothekerverbände)
ADL	Activities of daily living
ALT	Alanine aminotransferase
AML	Acute myeloid leukemia
ASCO	American Society of Clinical Oncology
ASS	Acetylsalicylic acid
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC	Area under the curve
BMI	Body mass index
CARG	Cancer and Aging Research Group
CCI	Charlson Comorbidity Index
CGA	Comprehensive geriatric assessment
CI	Confidence interval
CIRS	Cumulative illness rating scale
CNS	Central nervous system
CRASH	Chemotherapy Risk Assessment Scale for High-Age Patients
CTCAE	Common Terminology Criteria for Adverse Events
DDD	Defined daily dose
DIMDI	German Institute for Medical Documentation and Information (Deutsches Institut für Medizinische Dokumentation und Information)
DRP	Drug-related problem
ECOG	Eastern Cooperative Oncology Group
e.g.	For example (Latin ”exempli gratia”)
EORTC	European Organisation for Research and Treatment of Cancer
FORTA	Fit for the aged
GGT	Gamma-glutamyl transferase
GI	Gastrointestinal
GU	Genitourinary
HCT	Hydrochlorothiazide
IADL	Instrumental activities of daily living
i.e.	That is to say (Latin ”id est”)

IQR	Interquartile range
KPS	Karnofsky performance status
LDH	Lactate dehydrogenase
MAI	Medication Appropriateness Index
MMSE	Mini-Mental State Examination
MNA	Mini Nutritional Assessment
MOS	Medical Outcomes Study
n	Number of patients
NCI	National Cancer Institute
NCCN	National Comprehensive Cancer Network
nd	Not determinable
NHL	Non-Hodgkin lymphoma
NSAID	Nonsteroidal anti-inflammatory drugs
NSCLC	Non-small cell lung cancer
OR	Odds ratio
OTC	Over-the-counter
PIM	Potentially inadequate medication
PNP	Peripheral sensory neuropathy
PPI	Proton-pump inhibitors
PRO	Patient-reported outcomes
PRO-CTCAE	Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events
ROC	Receiver operating characteristic
ROC-AUC	Area under the receiver operating characteristic curve
rPDDI	Relevant potential drug-drug interactions
SD	Standard deviation
SIOG	International Society of Geriatric Oncology
SPSS	Statistical package for the social sciences
START	Screening Tool of Older Person’s Prescriptions
STOPP	Screening Tool to Alert Doctors to Right Treatment
ULN	Upper limit of normal
US	United States of America

Preliminary note

For the sake of clarity and to improve the readability the use of the female gender was largely avoided in the present study. The respective wording “he” is meant to include the female gender.

Furthermore, the author of this work was anxious to consider the copyright of all used texts, figures and data.

1 Introduction

1.1 Older cancer patients

Cancer is a disease of the elderly: The incidence of cancer substantially increases with higher age [1]. Due to demographic changes regarding a higher life expectancy and an aged population, the number of older cancer patients will continue to rise over the next years. This trend represents an increasing challenge since the therapy of older cancer patients is more complex than the therapy of younger cancer patients.

Pharmacokinetics and pharmacodynamics are altered at advanced age. Regarding pharmacokinetics, the most important physiological change affecting clinical routine is the decrease of renal function [2]. The glomerular filtration rate subsides since the renal blood flow, as well as renal mass, declines with age [3]. Renal impairment may entail higher plasma concentrations of drugs, possibly causing toxicity. Moreover, the hepatic elimination is decreased due to reduced hepatic volume, less hepatic blood flow and declined hepatic phase I metabolism (e.g. oxidative reactions) [3]. Apart from the elimination organs, likewise other physiological changes contribute to altered pharmacokinetics in older patients: The proportion of total body water decreases whereas the proportion of total body fat increases with age. This may lead to a reduced volume of distribution regarding hydrophilic drugs [2]. Delayed gastric emptying, a higher gastric pH value and reduced gastric motility [2], may alter pharmacokinetics as well, being especially important for oral antineoplastic therapies [3]. Concerning pharmacodynamics, older patients commonly show modified efficacy and tolerability of drugs [2]. Due to decreased homeostatic capacities and physiological reserves (e.g. bone marrow reserves), older patients are more vulnerable to therapy-related toxicity [3]. For example, older patients are more sensitive to adverse effects of central nervous system (CNS) drugs like benzodiazepines [2] and at higher risk for cardiomyopathy caused by anthracyclines [3].

Furthermore, older cancer patients often suffer from concomitant chronic conditions. A prospective study indicated that comorbidity was more common for older cancer patients (76%) than for younger cancer patients (51%) [4]. Vascular disorders were observed as most frequent comorbidity [4]. Comorbidity impairs cancer therapy due to a higher risk of

chemotherapy-related toxicity and hospitalization [5]. Moreover, it complicates therapeutic decisions due to missing evidence regarding this particular cohort and due to reduced life expectancy [5]. Since severe comorbidity comprises a competing risk for mortality [5], the benefit of cancer therapies might be limited in those patients. In general, different priorities might be important for older cancer patients: Since life expectancy is globally attenuated, short term quality of life and the ability to continue self-care in daily life might be of higher importance for older patients than a small advantage in survival [6].

Medication risks like polymedication are also common in older cancer patients, affecting cancer therapy as well. This aspect is further discussed in section 1.2.

Another issue comprises the paucity of evidence caused by the under-representation of elderly cancer patients in clinical trials [7]. Older patients are not enrolled as much as expected when considering cancer incidence in this age group [8]. Thus, there has recently been a call for pragmatic trials, facilitating the inclusion of older cancer patients. Pragmatic trials are characterized by a reduced burden on patients, broader eligibility criteria and a setting closer to everyday practice [9].

Patients at advanced age have shown to experience a higher toxicity risk during cancer therapy than younger patients [10, 11]. Nevertheless, older patients can benefit from chemotherapy as well as younger patients: Lichtman et al. observed that older women with breast cancer experienced similar efficacy as younger women – despite the higher risk of toxicity at higher age [10]. Muss et al. found similar results [11]. Nonetheless, many older patients are undertreated: Bouchardy et al. reported that half of older breast cancer patients were undertreated in their study, substantially worsening their survival [12]. Presumably, physicians do not expect older patients to cope with a more aggressive treatment.

However, older patients are a very heterogeneous population [2]. Aging does not simply depend on chronological age, counting the pure number of life years, but rather on the physiological age (“biological age”) which differs substantially: Some 80-year-old patients could be healthy and fit whereas others might be entirely frail, not being capable of self-care. Whether a patient will be able to go through a certain cancer therapy or not is thus difficult to judge, making therapy decisions highly complex.

Due to the heterogeneity of aging and the altered risk-benefit assessment, strategies for appropriate therapy individualization of older cancer patients are needed. The International Society of Geriatric Oncology (SIOG) recommends a comprehensive geriatric assessment (CGA) for individualizing cancer care in this population [13]. The geriatric assessment will be further discussed in section 1.3 of this thesis. Improvements in the treatment of older cancer patients are urgently warranted indeed: During the last years, the survival of older cancer patients increased slower than the survival of younger cancer patients, thereby leading to a larger survival gap between those cohorts [14].

1.2 Medication risks in older cancer patients

Drug-related problems are very common among older patients with cancer [15]. A drug-related problem is defined as an event during pharmacotherapy which interferes with a desired health outcome, for example an inappropriate timing of drug administration or overdosing [16]. A retrospective analysis showed that 90% of older cancer patients exhibited drug-related problems (DRP); in median three DRP per patient were detected [15]. The most frequent DRP were interactions (36.4%), adverse drug reactions (31.7%), and non-adherence (8.9%) [15]. Another study by Nightingale et al. found a DRP prevalence of 95% in older cancer patients, with a mean of three DRP per patient [17]. Those drug-related problems may have severe consequences for patients: An observational study in a Norwegian hospital indicated that about 4% of deaths among cancer patients occurred due to adverse drug events [18]. The National Comprehensive Cancer Network (NCCN) recommends a periodic medication review for older cancer patients [19]. The evaluation of medication is also regarded as an essential aspect of the geriatric assessment (see section 1.3), being recommended by the SIOG [13, 20].

There are in particular three aspects of medication risks which are usually highlighted for older cancer patients: Polymedication, potentially inadequate medication, and drug-drug interactions.

1.2.1 Polymedication

Polymedication is usually defined as an intake of five drugs or more [21]. Different definitions of polymedication exist but the cut-off "≥ 5" is commonly used and has shown to be associated with adverse outcome in the elderly [22]. In general, older patients take more drugs than younger patients: In Germany, 55% of the defined daily doses (DDD) of drugs paid by the statutory health insurance was taken by patients ≥ 65 years in 2015 – although this age group only comprised 22% of the total population [23]. About one third of patients ≥ 65 years was found to be exposed to polymedication [24]. This high prevalence of polymedication was also detected in older patients with cancer: A study with 385 older cancer patients observed a prevalence of 57% for polymedication (≥ 5 drugs), with on average about six drugs per patient being taken [25]. Another study indicated that 80% of 117 older patients used five or more drugs prior to the start of cancer treatment [26]. Being associated with adverse outcomes like increased mortality or hospitalization, polymedication is of concern for the older population in general [25]. In older cancer patients, adverse outcomes were reported as well: Polymedication was associated with frailty and decreased physical function [25, 27]. Moreover, it was related to a 6-fold increased odds of experiencing severe chemotherapy-related toxicity in older metastatic breast cancer patients [28].

1.2.2 Potentially inadequate medication

When judging medication quality in older cancer patients, it is not only important to consider *how many* drugs, but also *which* drugs are being taken. There might be good reasons to prescribe many drugs to an older person. Hence, instead of only discussing the number of drugs, "appropriate polymedication" should rather be in focus, recognizing that patients can benefit from numerous medicines if they are chosen based on evidence and the clinical context [29]. The appropriateness of medication use is considered in this study via the screening of potentially inadequate medication (PIM). At higher age, some drugs may lead to an increased mortality risk, to a higher risk of falls or to more adverse reactions [21]. For these drugs, the risks may outweigh benefits in older patients [21]. Those PIM drugs for elderly can be determined in different ways: On the one hand, explicit PIM lists were developed, which are drug-oriented, explicitly stating specific drugs as being potentially inadequate [30]. On the other hand, implicit tools are available, which are judgment-based and patient-specific [30].

Examples for explicit PIM lists comprise the Beers list [31] (mainly developed for the US), the PRISCUS list [32] (tailored to Germany), and the EU(7)-PIM list [33] (adapted to Europe). Also, FORTA (fit for the aged) [34] and START/STOPP (Screening Tool of Older Person's Prescriptions/Screening Tool to Alert Doctors to Right Treatment) criteria [35] are interesting to mention, since those additionally list drugs being useful in older patients ("positive list"). An example for implicit PIM lists comprises the Medication Appropriateness Index (MAI) which evaluates the medication regarding ten criteria (for instance: "Is the dosage correct?" or "Is there unnecessary duplication with other drug(s)?") [36]. An advantage of the explicit tools comprises their applicability without clinical judgment. However, they do not take the individual patient into consideration for assessing appropriateness and need to be updated constantly [30]. In contrast, implicit lists individualize the assessment for each patient. However, the judgment depends on the expertise of the user and is time-consuming [30]. Depending on the instrument, different PIM prevalence was found in literature. A study with 160 older patients receiving parenteral cancer therapy in an ambulatory clinic found that 48.1% used at least one PIM (2015 Beers criteria) [37]. Another study detected a PIM prevalence of 38% according to the STOPP criteria [38]. PIM have shown to be associated with adverse outcomes in the general older population. Reich et al. found PIM to be associated with the hospitalization of patients [39]. In another study, PIM use was related to an increase of adverse drug events [40]. For older patients with cancer, only few studies exist regarding the association of PIM and adverse outcomes. So far, no association was found in respective studies [22].

1.2.3 Drug-drug interactions

Since older cancer patients frequently experience polymedication, they are also at an increased risk of drug-drug interactions [41]: Riechelmann et al. found an increasing number of drugs to be associated with a higher risk of potential drug-drug interactions in the general population of cancer patients [42]. Potential drug-drug interactions are frequent among older cancer patients: Yeoh et al. detected potential drug-drug interactions as the most frequent drug-related problem (36.4%) in older patients receiving outpatient chemotherapy [15]. In another study, 75.4% of older cancer patients experienced potential drug-drug interactions during therapy [43]. Clinical consequences of those interactions might be serious: A study

reported that drug-drug interactions caused unplanned hospitalizations of cancer patients in about 2% of cases [44].

1.3 Geriatric assessment

The functional abilities of older cancer patients are heterogeneous [2].Therefore, it is essential to individualize therapy. Clinical judgment, chronologic age, or performance status are usually used for this in daily routine [45]. Commonly used scores are the Eastern Cooperative Oncology Group (ECOG) performance status or the Karnofsky performance status; see Table 1-1. However, those approaches were not found to be adequate for therapy individualization. Clinical judgment was reported to be less effective in selecting older patients for aggressive chemotherapy than a comprehensive geriatric assessment (CGA) [45]. Chronologic age was observed to be less predictive for chemotherapy toxicity than overall health [46]. The performance status did not show adequate abilities either: Patients who scored normally on the Karnofsky performance status nevertheless exhibited substantial impairments identified by a geriatric assessment [47].

Thus, for individualizing cancer care, a CGA is recommended for older cancer patients ≥ 70 years by the SIOG [13]. A CGA is a multidimensional, interdisciplinary evaluation of older patients to identify care needs regarding e.g. physical health, psychosocial, or functional capabilities [13, 48]. Based on this, a treatment plan can be developed [13, 48]. In geriatrics, a CGA is based on the following four domains: physical health, functional status, psychological health, and socioenvironmental factors [48]. In geriatric oncology, the domains of a CGA have not been well defined and not all studies conducted a complete CGA [48]. The SIOG recommends evaluating the following domains for geriatric assessments in older cancer patients: functional status, fatigue, comorbidity, cognition, mental health status, social status and support, nutrition, and geriatric syndromes [20]. Geriatric syndromes include for example delirium, dementia, falls, incontinence, but also polymedication [20]. Different instruments for assessing those domains exist. However, the SIOG could not recommend one instrument above the other, advising to choose the specific instrument according to local preferences or resources [20]. An overview of essential domains of the geriatric assessment is given in Table

1-2. Furthermore, the NCCN recommends a periodic medication review in older cancer patients [19].

Table 1-1 *Comparison of the ECOG* [49] *and Karnofsky performance status* [50], according to [51]; *ECOG, Eastern Cooperative Oncology Group*

Score	ECOG performance status	Score	Karnofsky performance status
0	Fully active, able to carry on all pre-disease performance without restriction	100	Normal, no complaints; no evidence of disease
		90	Able to carry on normal activity; minor signs or symptoms of disease
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature	80	Normal activity with effort, some signs or symptoms of disease
		70	Cares for self but unable to carry on normal activity or to do active work
2	Ambulatory and capable of all self-care but unable to carry out any work activities; up and about more than 50% of waking hours	60	Requires occasional assistance but is able to care for most of personal needs
		50	Requires considerable assistance and frequent medical care
3	Capable of only limited self-care; confined to bed or chair more than 50% of waking hours	40	Disabled; requires special care and assistance
		30	Severely disabled; hospitalization is indicated although death not imminent
4	Completely disabled; cannot carry on any self-care; totally confined to bed or chair	20	Very ill; hospitalization and active supportive care necessary
		10	Moribund
5	Dead	0	Dead

Table 1-2 Domains of a geriatric assessment in the older cancer patients including examples for respective instruments; modified according to [20]*; MOS, Medical Outcomes Study; START, Screening Tool to Alert Doctors to Right Treatment; STOPP, Screening Tool of Older Person's Prescriptions*

Domain	**Instruments (examples)**
Social status	Questions on living situation, marital status, financial resources MOS social support survey
Comorbidity	Charlson Comorbidity Index (CCI) Cumulative illness rating scale (CIRS)
Functional status	Timed get up and go Hand grip strength Activities of daily living (ADL) Instrumental activities of daily living (IADL) Barthel index
Cognition	Mini-Mental State Examination (MMSE) Clock-drawing test
Depression	Geriatric depression scale Hospital anxiety and depression scale
Nutrition	Body mass index (BMI) Mini Nutritional Assessment (MNA)
Medication	Beers criteria STOPP and START criteria
Geriatric syndromes	Polymedication Falls Constipation Pressure ulcus

1.3.1 Usefulness

The geriatric assessment was developed in geriatrics and has shown to be effective in the general older population: A CGA was found to increase survival and increase the probability of staying at home up to 12 months for older patients admitted to hospital [52].

In older cancer patients, the geriatric assessment was found to identify age-related problems not being revealed by routine diagnostics: Kenis et al. observed that geriatric interventions

were initiated in 25.7% of patients based on a geriatric assessment [53]. Furthermore, the geriatric assessment affected therapy decisions in older cancer patients and thus demonstrated an impact on clinical routine: A systematic review reported that 21% to 53% of treatment regimens were altered after a geriatric assessment, in particular due to the functional or nutritional status [54]. In line with this, another review observed that cancer treatment plans were changed in a median of 28% of patients following a geriatric assessment [55].

Moreover, the results from geriatric assessments were associated with adverse outcomes in older cancer patients: A systematic review by Versteeg et al. found that different geriatric assessment domains were associated with mortality in older cancer patients [54]. Nutritional status predicted mortality in all analyzed studies. No consistent parameters were found for predicting toxicity [54]. A systematic review by Caillet et al. reported that each geriatric assessment domain was associated in at least one study with therapy-related toxicity and survival. In particular, malnutrition, comorbidity, and functional decline seemed to be predictive [56]. A systematic review by Hamaker et al. found little consistency regarding the value of geriatric assessment parameters for predicting outcomes in older cancer patients. For instance, frailty was associated with toxicity, whereas cognitive function and ADL were associated with completion of cancer therapy [57].

However, for determining the clinical impact of a geriatric assessment, it is also important to investigate if it may effectively improve outcomes of patients. Only few studies analyzed the impact of a geriatric assessment on outcomes. However, these studies showed mixed results. Corre et al. assessed the benefit of a geriatric assessment on outcomes in a randomized multi-centric study, using the CGA versus a standard approach (based on ECOG performance status and age) for assigning patients to different therapy strategies [58]. The primary endpoint, treatment failure free survival, did not show an effect of the CGA. However, the CGA could significantly reduce treatment toxicity [58]. In a prospective cohort comparison study, Kalsi et al. found that patients with CGA were significantly more likely to complete the planned therapy regimen and underwent less therapy modifications [59]. Treatment toxicity was reduced as well but did not reach statistical significance [59]. A randomized pilot study by Magnuson et al. did not observe an improvement of outcomes (for example chemotherapy-related grade 3-5 toxicity, hospitalization, or early treatment discontinuation) in patients with

geriatric assessment compared to standard care [60]. This might be explained by the low implementation rate of geriatric assessment recommendations by the primary oncologist in this study [60].

1.3.2 Implementation into daily routine

The SIOG summarizes that a geriatric assessment may be helpful in daily practice for detecting geriatric deficits of patients, for adapting treatment choice, and for predicting severe therapy-associated toxicity, as well as overall survival [20]. The main goal of a geriatric assessment should be to trigger interventions for geriatric deficits and to guide individualized treatment decisions [20]. The geriatric assessment should support decision-making. However, it is too early to deny or assign an oncological therapy only on grounds of a geriatric assessment [61].

Despite being recommended, a geriatric assessment is not always feasible in daily routine, being very time-consuming [62]. However, efforts have been made during the last years to implement the geriatric assessment into daily routine using interdisciplinary concepts. For instance, Schmidt et al. implemented an interdisciplinary care concept for older cancer patients in a German university hospital, including results of a geriatric assessment and patient-reported quality of life [63].

For facilitating the implementation of the time-consuming geriatric assessment, several short frailty screening tools were developed for pre-selecting patients in need of a CGA. However, a systematic review by Hamaker et al. concluded that none of the available screening tools indicated sufficient discriminative abilities for selecting patients requiring a full CGA [64]. Nevertheless, due to time-savings, the SIOG concluded that screening tools might be used for pre-selecting patients in busy clinical routine [65]. No screening tool could be recommended above another by SIOG [65].

1.4 Onco-geriatric scores

Toxicity is an important aspect for therapy decisions in older cancer patients [66]. Prediction of toxicity may help oncologists as well as patients in the decision-making process regarding cancer therapy [67]. Short tools combining geriatric assessment parameters with oncologic

parameters were developed for individualized prediction of chemotherapy-related toxicity: the CARG (Cancer and Aging Research Group) and the CRASH (Chemotherapy Risk Assessment Scale for High-Age Patients) score [67, 68].

An advantage of the onco-geriatric scores comprises the inclusion of oncological items for toxicity prediction [67, 68]. The geriatric assessment is recommended for therapy individualization [13] and has shown to be associated with therapy-related toxicity [56]; however, also oncological parameters (for example the therapy regimen [69]) play a role for the toxicity risk. In contrast to the onco-geriatric scores, other frailty screening tools focus on geriatric assessment items only [65]. Furthermore, the onco-geriatric scores offer an easier interpretation of the toxicity risk. A geriatric assessment itself does not inform about risk probabilities, whereas the scores yield a certain toxicity risk category [67, 68]. This facilitates the applicability of results to therapy decisions. Moreover, the onco-geriatric scores are short: Instead of incorporating the full geriatric assessment, only those domains predictive for toxicity were extracted [67, 68]. This also facilitates their use in clinical routine.

The American Society of Clinical Oncology (ASCO) guideline for Geriatric Oncology recommends either the CARG or the CRASH score for predicting toxicity [70]. Although both scores are recommended in literature, it is unclear which score should be preferred [70]. Whereas the CARG score is a short prediction tool incorporating only few simple questions, the CRASH score is more time-consuming, including full geriatric assessment instruments (IADL, MMSE, MNA) [67, 68]. However, the CRASH score could possibly give a more detailed prediction, differentiating between hematologic and nonhematologic toxicity as well as between four categories while the CARG score only predicts a general toxicity incidence regarding three risk categories [67, 68]. In a review by Almodovar et al., a panel of six experts judged both scores as feasible tools in non-small cell lung cancer (NSCLC) treatment but considered CARG as the first option in clinical routine due to its ease of use [62]. However, no study has compared the predictive performance of the CARG and CRASH score so far.

The CARG score was evaluated for different tumor entities, demonstrating mixed results: In lung cancer patients, toxicity incidence increased significantly with higher CARG risk categories, suggesting a predictive value of the CARG score [71]. For prostate cancer patients, the CARG score could not demonstrate a predictive value for therapy-related toxicity [72]. However, this study was limited by a relatively small sample size (46 patients) [72]. A recent

study investigated the CARG score in 126 patients with solid tumors, not finding a sufficient predictive value of the CARG score [73]. In contrast to that, another study with 58 older patients with solid tumors found that patients with a CARG score ≥ 10 experienced more often toxicity than patients with a CARG score of < 10 [74]. The CRASH score has not been assessed in further studies so far.

Due to the conflicting results and lacking data, further evaluation and comparison of the onco-geriatric scores is urgently needed to support the implementation of those promising tools in daily routine. The onco-geriatric scores may be of significance in different fields. In clinical routine, the onco-geriatric scores could allow for a more detailed weighting of risks and benefits of therapies. This might in consequence also reduce under- and overtreatment of older cancer patients. Furthermore, clinical trials may use these scores for risk stratification.

2 Aim

The overarching goal of this work consisted in optimizing the treatment of older cancer patients. This was pursued by (I) evaluating onco-geriatric scores (the CARG and the CRASH score) and by (II) assessing medication risks in older cancer patients.

Regarding the onco-geriatric scores, the primary aim was to compare the CARG and CRASH score concerning (I) the agreement of predictions and (II) the predictive performance regarding toxicity risks. The secondary aim consisted in comparing the score predictions with physicians' judgment and with other commonly used predictors of toxicity. For investigating additional applications of the scores, exploratory analyses were conducted concerning the predictive value for time-related predictions, patient-reported symptom burden, and alterations of planned treatment.

Moreover, this thesis aimed at analyzing the medication risks of older cancer patients. Those were investigated before and after start of cancer therapy regarding three aspects: polymedication, potentially inadequate medication, and drug-drug interactions.

3 Methods

In the following section, the methods regarding the evaluation of onco-geriatric scores and the analysis of medication risks in older cancer patients are described.

3.1 Evaluation of onco-geriatric scores

Onco-geriatric scores (the CARG and the CRASH score) were developed for predicting the occurrence of severe therapy-associated toxicity in older cancer patients. This study evaluated and compared the predictive performance of both scores for determining which one is favorable to use in a clinical routine setting.

3.1.1 Study design

The study was a prospective, single-center, observational study. All study patients were not treated according to a pre-specified study protocol but solely according to clinical routine. The physicians' decision on treatment, diagnosis, and monitoring was not affected by the study protocol. Legally, this study is therefore classified as a non-interventional trial according to §4 of the German drug law (Arzneimittelgesetz) [75].

A positive vote of the ethics committee of the Faculty of Medicine of Bonn University was granted for this study (consecutive number 302/15).

This study was conducted in cooperation with the Department of Geriatrics and Neurology (Prof. A. Jacobs) as well as the Department of Oncology and Hematology (Prof. Y.-D. Ko) at the Johanniter Hospital Bonn. The Johanniter Hospital Bonn features 364 beds and is part of the cancer center of the Bonn/Rhein-Sieg area, certified by the German Cancer Society (Deutsche Krebsgesellschaft e.V.) [76, 77]. It offers inpatient oncological care as well as an outpatient oncological clinic.

3.1.2 Pilot study

This evaluation study was set up based on a previous pilot study by our research group (positive vote of the ethics committee of the Faculty of Medicine of Bonn University, consecutive number 011/15) conducted in the Johanniter Hospital Bonn between March and June 2015. The prospective, observational pilot study aimed at testing the feasibility of the CARG and the CRASH score performance in a clinical routine setting (see details of the CARG and the CRASH score in 3.1.5.1 and 3.1.5.2). The CARG and the CRASH score were performed in 20 patients $\geq$ 70 years who experienced a malignancy or condition requiring treatment with antineoplastic agents. Patients must not have started systemic cancer treatment yet. Eligibility criteria were consistent with the subsequent evaluation study, except regarding the performance of the systemic cancer treatment. In the pilot study, a systemic cancer treatment had to be indicated as standard therapy but, contrary to the evaluation study, the actual performance of the treatment was not required. If no systemic treatment was conducted, score items regarding cancer treatment were calculated using the standard therapy. Parts of the pilot study were published in the Master thesis of Monique Theissen Mendel [78]. Data of the pilot study were partly included into the analysis of medication risks (see section 3.2).

3.1.3 Course of the study

This study considered the design and eligibility criteria of the original development studies of the CARG and the CRASH score where possible. Patients were recruited at the oncology and internal medicine inpatient wards of Johanniter Hospital Bonn. In a patient interview, the geriatric assessment items for the CARG and CRASH score (see sections 3.1.5.1 and 3.1.5.2) as well as a baseline assessment of the symptom burden (PRO-CTCAE, see section 3.1.6.2) were captured. Additional laboratory data for the scores (see sections 3.1.5.1 and 3.1.5.2) and patient characteristics (see section 3.1.4.4) were collected from medical records. For the CARG and CRASH score calculation, either laboratory data of the day of inclusion or the most recent measurements were used. Before start of systemic cancer therapy, the treating physician was asked about the planned cancer treatment as well as the clinical judgment in terms of toxicity risk (physicians' judgment; see section 3.1.5.3). Physicians were blinded to the score results; hence, the results did not influence either diagnosis, treatment decision, or monitoring.

As follow-up, toxicity during therapy course (CTCAE, see section 3.1.6.1) and symptom burden of patients (PRO-CTCAE, see section 3.1.6.2) were captured until the end of the planned therapy or for a maximum of six cycles. The toxicity during therapy course and possible alterations of planned treatment (therapy discontinuations, changes, delays, or reductions) were retrospectively recorded from medical records. Patients were contacted for the assessment of symptom burden one to two weeks after start of cycle. Follow-up data, including doctor letters and laboratory data, were collected in the Johanniter Hospital Bonn as well as in surrounding oncology practices if patients continued treatment there. The data collection, comprising the performance of the onco-geriatric scores and capturing of follow-up data, was delivered by the author of this thesis (in the following referred to as "researcher"). Details regarding the course of the study are given in Figure 3-1.

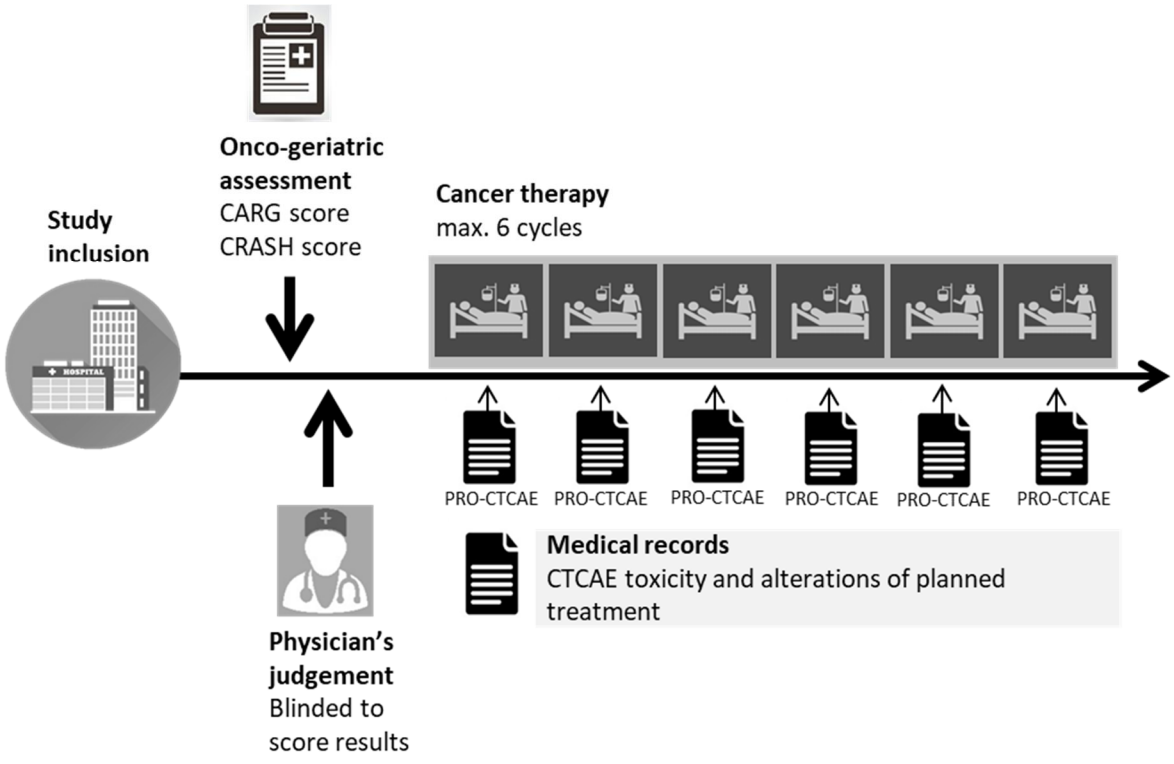


Figure 3-1 Course of the study and outcome measurements; PRO-CTCAE, Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events; CTCAE, Common Terminology Criteria for Adverse Events

3.1.4 Patients

The patient cohort consisted of older cancer patients starting a first-line systemic cancer treatment in an inpatient clinical routine setting.

3.1.4.1 Inclusion and exclusion criteria

Following inclusion and exclusion criteria were defined:

Inclusion criteria

- Diagnosis of any malignancy (solid tumor or hematologic malignancy)
- Age ≥ 70 years
- Understanding German
- Systemic cancer treatment indicated (chemotherapy, targeted therapy, or immunotherapy)
- Written informed consent

Exclusion criteria

- Cognitive impairment which prevents understanding of the course or purpose of the study (exclusion of patients with a Mini-Mental State Examination < 20)
- Systemic cancer treatment had already started

Broad inclusion criteria and only limited exclusion criteria were defined because the onco-geriatric scores were designed to comprise a broad prediction scope across different cancer types and cancer therapies. Moreover, a heterogeneous study population better represents the real population of the hospital. The original development studies of the CARG and CRASH score had different inclusion criteria, see sections 3.1.5.1 and 3.1.5.2. For this evaluation study, eligibility criteria were selected based on both development studies.

3.1.4.2 Recruitment

For recruitment, patients at the oncology and internal medicine wards of the Johanniter Hospital Bonn were identified by the treating physicians and invited to participate. All patients fulfilling the defined inclusion and exclusion criteria were enrolled in the study. Patients were

informed about the study details, received a patient information brochure and signed a written informed consent. The patient information brochure and the informed consent form are presented in Appendix A. Patient data were pseudonymized by assigning a random number to each patient. After the end of the study, data were anonymized.

3.1.4.3 Sample size determination

In general, it is recommended to include 100 events and 100 nonevents for an external validation of prediction models [79]. However, since data about the incidence of severe toxicity in this patient cohort was missing, a sample size calculation was not possible. Based on the experiences from the pilot study, a recruitment rate of approximately six patients per month was estimated and 120 patients were sought to be included.

3.1.4.4 Documentation

The following patient characteristics were documented: demographic data, renal function (Cockcroft-Gault), comorbidity, diagnosis of cancer, cancer therapy regimen, and medication of the patient. For comorbidity, the Charlson Comorbidity Index (CCI) [80] in its original version was calculated since this tool was found to be reliable in older cancer patients and is the most widely used comorbidity tool [81, 82]. The cancer diagnosis was neglected for the CCI calculation since the focus of this index was to describe the additional conditions beside the primary cancer diagnosis. All required laboratory data were measured in routine care; thus, no additional blood draws were necessary.

3.1.5 Risk assessment

The risk of therapy-associated toxicity was assessed by the CARG score, the CRASH score, and physicians' judgment. For details regarding study material of the CARG and CRASH score see Appendix B.

3.1.5.1 The CARG score

The CARG score was developed by Hurria et al. in 2011 and externally validated in 2016 [67, 83]. It was developed in a prospective, multicentric study with 500 patients ≥ 65 years [67]. All patients experienced solid tumors and were scheduled to start a new chemotherapy regimen in an outpatient oncology practice. The mean age was 73 years and most patients experienced lung cancer (29%) or gastrointestinal cancer (27%) during the development study. The predicted outcome was therapy-associated severe toxicity, defined as grade 3-5 toxicity according to the Common Terminology Criteria for Adverse Events (CTCAE). Two physicians reviewed chemotherapy courses for recording therapy-related toxicity. Toxicity was documented at each clinical encounter. Severe toxicity regarding blood values was only considered if it had occurred on the day of scheduled chemotherapy or at emergency visits. During the development study, 53% of patients experienced severe toxicity. Various geriatric assessment parameters, laboratory data, and tumor- or treatment-related oncological variables were investigated for associations with severe toxicity. For the CARG score, eleven of those variables were selected, including five geriatric assessment items, three clinical items, and three oncological items. The items included in the score are presented in Table 3-1, together with the assigned score points [67].

The CARG score ranges from 0 to 23 points. Higher points indicate a higher risk. In this evaluation study, the original cut-offs for the score categories were used: Score results ranging from 0 to 5 points were assigned to predict a low risk of therapy-associated toxicity, 6 to 9 points a mid risk, and ≥ 10 points a high risk. In the development cohort, the proportion of patients with low, mid, and high toxicity risk were 30%, 52%, and 83%, respectively [67].

In this evaluation study, the CARG score online calculator was used for computing the toxicity score (http://www.mycarg.org, last accessed 2019 September 28). In order to determine if a regimen was categorized as “polytherapy or monotherapy”, all antineoplastic agents were counted, including targeted therapies or immunotherapies. The dosage was regarded as “reduced or standard” according to the treating physician, the summary of product characteristics, or therapy guidelines, as appropriate. The tumor entities “gastrointestinal or genitourinary” were classified following the definition of the National Cancer Institute [84]. In this evaluation study, all geriatric assessment items were captured in a patient interview; the clinical and oncological variables were extracted from medical records. The geriatric

assessment questions were translated into German based on the original English CARG version [67].

Table 3-1 Items of the CARG score with the assigned score points [67]

Area	Item	Value	Score points
Patient characteristics	Age	≥ 72 years	2
		< 72 years	0
Tumor characteristics	Tumor entity	Gastrointestinal or genitourinary	2
		Other	0
Therapy regimen	Dosage	Standard	2
		Reduced	0
	Number of chemotherapeutic agents	Polytherapy	2
		Monotherapy	0
Laboratory items	Hemoglobin	< 10 g/dL (female), < 11 g/dL (male)	3
		≥ 10 g/dL (female), ≥ 11 g/dL (male)	0
	Creatinine clearance (Jeliffe) [85]	< 34 mL/min	3
		≥ 34 mL/min	0
Hearing	Hearing ability	Fair/worse	2
		Good/excellent	0
Functional status	Number of falls (in the last 6 months)	≥ 1	3
		0	0
	Help in taking medications	Requires assistance	1
		No assistance	0
	Ability of walking one block	Limited	2
		Not limited	0
	Decreased social activity due to physical/emotional health	Some, most, all of the time	1
		A little, or none of the time	0

In the development study, the CARG score resulted in an area under the receiver operating characteristic curve (ROC-AUC) of 0.72 [67], which implies rather moderate discriminative

abilities of the score (see section 3.1.7 for details on ROC analyses). In the validation study with 250 patients, a ROC-AUC of 0.65 was reached which was not statistically different from the ROC-AUC in the development cohort [83]. In both studies, the CARG score exhibited a better toxicity prediction than the Karnofsky performance status, a commonly used predictor for toxicity [50, 67, 83].

Instead of using complete instruments from the geriatric assessment, the CARG score only uses single questions being extracted from those instruments (e.g. from the IADL only one question was incorporated: “help with medication intake”) [67]. Thus, the CARG score offers a quick estimation of the toxicity risk and can be completed in < 5 min [70] . However, only three different risk categories are used and it does not offer a risk prediction for different types of toxicity but only for overall toxicity [67].

3.1.5.2 The CRASH score

The CRASH score was developed by Extermann et al. in 2012 [68]. The prospective, multicentric study included 562 patients $\geq$ 70 years with solid and hematologic tumors, starting a new chemotherapy regimen [68]. The score was concomitantly developed and validated in that study: The patient cohort was randomly split in a derivation and a validation cohort (ratio 2:1). The median age of the derivation cohort was 76 years and the most frequent tumors comprised lung cancer (21.5%), breast cancer (21.5%), and non-Hodgkin lymphoma (14.2%). As outcome, the occurrence of severe toxicity during therapy course was analyzed. “Severe toxicity” was defined as CTCAE grade 3-4 nonhematologic or grade 4 hematologic toxicity. Toxicity was captured via medical evaluation at start of each cycle and at the end of treatment as well as via screening of medical records. Weekly complete blood counts were considered. The chemotherapy follow-up was ended after a maximum of 6 months. During the development study, severe toxicity occurred in 64% of patients [68]. Different clinical variables, laboratory data, geriatric assessment instruments, and cancer-specific variables were investigated for toxicity prediction. The derived score was divided into three subscores predicting overall (combined CRASH score), hematologic (hematologic CRASH score), and nonhematologic toxicity (nonhematologic CRASH score). The combined score was constructed by combining the items of the hematologic and nonhematologic score into one score. The

items included in the CRASH score are presented in Table 3-2, together with the assigned score points [68].

Table 3-2 Items of the CRASH score and the assigned score points regarding the three different subscores hematologic, nonhematologic, and overall toxicity [68]*; ULN, upper limit of normal*

CRASH subscore	Area	Item	Item value	Score points
Hematologic	Cancer therapy	MAX2 index [69] (Chemotherapy toxicity index)	> 0.57	2
			0.45-0.57	1
			0-0.44	0
	Laboratory item	Lactate dehydrogenase (LDH)	> 0.74 x ULN	2
			0-0.74 x ULN	0
	Clinical item	Diastolic blood pressure	> 72 mmHg	1
			≤ 72 mmHg	0
	Functional status	Instrumental activities of daily living (IADL) [86]	10-25	1
			26-29	0
Nonhematologic	Cancer therapy	MAX2 index [69] (Chemotherapy toxicity index)	> 0.57	2
			0.45-0.57	1
			0-0.44	0
	Functional status	Eastern Cooperative Oncology Group (ECOG) performance status [49]	3-4	2
			1-2	1
			0	0
	Cognitive function	Mini-Mental State Examination (MMSE) [87]	< 30	2
			30	0
	Nutritional status	Mini Nutritional Assessment (MNA) [88]	< 28	2
			28-30	0
Combined	Addition of score points of hematologic and nonhematologic scores with MAX2 only counting once			

To adjust the score to the general toxicity of a chemotherapy regimen, the MAX2 index was used. This index has been previously developed by Extermann et al. and specifies the general per-patient toxicity risk of a chemotherapy regimen [69]. The index is derived from clinical trial

data by computing the average of the highest frequency of nonhematologic grade 3-4 toxicity and hematologic grade 4 toxicity of a regimen. Typically, three published studies with at least 20 patients are include in the calculation [89]. Since taking into account the maximal frequency of both toxicity types, the index is called “MAX2” [69]. In the publication of the CRASH score, an overview of the MAX2 classifications was given for several typical chemotherapy regimens [68]. However, not all regimens were included in this list. In this case, an extended MAX2 list was used, which Extermann had provided to the author of this thesis on request. If the regimen was not listed there either, the regimen was classified by analogy, as recommended in the publication of the CRASH score [68]. The classification for the regimens not listed in the original publication of the CRASH score is presented in Appendix B. The German versions of the Mini-Mental State Examination (MMSE) [90] and the Mini Nutritional Assessment (MNA) [91] were used; the IADL [86] was translated into German based on the English original version. The geriatric assessment was performed by the researcher via a patient interview; clinical and laboratory data were recorded from medical records.

The combined, hematologic, and nonhematologic CRASH scores range from 0 to 12, 0 to 6, and 0 to 8 score points, respectively [68]. With increasing risk score, the risk of toxicity increases. For calculating the CRASH score, an online calculator tool was used (https://www.moffitt.org, last accessed 2019 September 28). In this evaluation study, the original cut-offs of the score categories were applied for analysis. For the corresponding categories of the score results and the respective toxicity incidence in the derivation cohort see Table 3-3 [68].

Table 3-3 Risk categories of the subgroups of the CRASH score, together with the observed proportion of toxicity incidence in the derivation cohort [68, 92]

Risk category	**Combined**		**Hematologic**		**Nonhematologic**	
Low	0-3:	50%	0-1:	7%	0-2:	33%
Mid-Low	4-6:	58%	2-3:	23%	3-4:	46%
Mid-High	7-9:	77%	4-5:	54%	5-6:	67%
High	> 9:	79%	> 5:	100%	> 6:	93%

The score subgroups where hematologic and nonhematologic toxicity were evaluated separately differentiated better than the combined score [68]. The CRASH score reached a

ROC-AUC in the same range as the CARG score: In the derivation cohort, the combined, hematologic, and nonhematologic CRASH score yielded a ROC-AUC of 0.65, 0.76, and 0.66, respectively. In the independent sample validation, lower ROC-AUC values for the combined, hematologic, and nonhematologic CRASH score were observed: 0.64, 0.65, and 0.62, respectively [68].

Compared to the CARG score, the CRASH score is more time-consuming, requiring approximately 20-30 min to complete [70]. However, the CRASH score could be easily integrated into a CGA, already fully including various geriatric assessment instruments (IADL, MMSE, MNA) [70]. In contrast to the CARG score, the CRASH score offers a differentiation between the types of toxicity and a more detailed category distinction in four risk categories [68].

3.1.5.3 Physicians' judgment

Before the start of cancer therapy, the treating physicians were asked to estimate their patient's individual toxicity risk during therapy course. The physicians should specify risk estimates for overall, hematologic, and nonhematologic severe toxicity, classified in the categories low, mid, or high. No detailed probability in percentage was requested because physicians are not trained for this detailed risk prediction and, therefore, would probably not be capable of giving such an exact estimate. All physicians were blinded to the score results; thus, the risk estimation was not influenced by the onco-geriatric assessment.

3.1.6 Outcome measurement

The primary endpoint was defined as severe toxicity during therapy course according to the Common Terminology Criteria for Adverse Events (CTCAE) [93]. Secondary endpoints comprised severe symptom burden of patients according to the Patient-Reported Outcomes version of the Terminology Criteria for Adverse Events (PRO-CTCAE) [94] and alterations of the planned treatment during therapy course.

Baseline values of the outcome parameters were captured. Follow-up was pursued until the end of therapy or until a maximum of six therapy cycles was reached. Since the planned

treatment itself was also included in the score risk prediction, the follow-up was not pursued any longer if patients completely changed the planned treatment regimen or experienced dose reductions ≥ 50%. If minor changes of the treatment regimen occurred (e.g. delay of treatment for a few days), the follow-up was continued since, in general, this did not substantially alter score predictions. All patients with at least one cycle of follow-up were included into the outcome analysis in order to use the maximum information available. The duration of each cycle was retrieved from the cancer therapy plan of each patient. An overview of the outcome measurements during the study is illustrated in Figure 3-1 (see above).

3.1.6.1 Toxicity

The toxicity during therapy course was captured from medical records according to the Common Terminology Criteria for Adverse Events (CTCAE), version 4.03 [93]. The CTCAE were developed by the US National Cancer Institute as standardized terminology for reporting adverse events. The severity of adverse events is described by CTCAE grades, ranging from 1 (mild) to 5 (death); for general definitions of grades see Table 3-4 [93].

Table 3-4 *General definitions of the CTCAE grades according to the National Cancer Institute* [93]*; grades defined as "severe toxicity" for this evaluation study are shown in bold; semi-colon signifies "or"*

Grade 1	Mild*;* asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated
Grade 2	Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living
Grade 3	**Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care activities of daily living**
Grade 4	**Life-threatening consequences; urgent intervention indicated**
Grade 5	**Death related to adverse events**

In this evaluation study, "severe toxicity" was defined as CTCAE grade ≥ 3. That definition is equal to the definition in the CARG score development study [67]. However, it differs from the

CRASH score development study where grade 3-4 nonhematologic toxicity or grade 4 hematologic toxicity was considered as “severe toxicity” [68].

An overview of all recorded types of toxicity is presented in Table 3-5. These were selected based on the toxicity captured during the CARG and CRASH score development studies [67, 68]. Additional items relevant for targeted or immunotherapy were included (e.g. skin reactions, hypertension). The selection of additional items was undertaken by literature review and the discussion of relevant symptoms with an experienced oncologist. The toxicity documentation form being used for data collection is displayed in Appendix C.

Table 3-5 Recorded types of toxicity during therapy course

Hematologic toxicity	Anemia, febrile neutropenia, neutropenia, leukopenia, thrombopenia
Nonhematologic toxicity	Acute coronary syndrome, atrial flutter, heart failure, hypertension, thromboembolic event, aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyl-transpeptidase (GGT), bilirubin, creatinine, proteinuria, dyspnea, erythroderma, urticaria, palmar-plantar erythrodysesthesia syndrome, gastrointestinal bleeding, infections, dehydration, hyponatremia, hypokalemia, dysphagia, dry mouth, mucositis, pain, anorexia, constipation, diarrhea, nausea, vomiting, insomnia, fatigue, peripheral sensory neuropathy

All medical records of the patients at Johanniter Hospital Bonn were retrospectively screened for toxicity. The highest CTCAE grade during each cycle was documented for each toxicity type. If no information was given about the severity of toxicity, CTCAE grade 1 was assumed. Further causality assessment was not performed, all detected toxicity was included. If a patient continued treatment in an outpatient clinic or oncology practice not located at the Johanniter Hospital Bonn, the clinic or practice was contacted for medical reports and laboratory data. Regular, weekly blood controls were considered for the toxicity evaluation. If these blood controls were not conducted at the Johanniter Hospital Bonn but e.g. at the general practitioner, the physician in charge was contacted for those data. Hence, the nadir of blood parameters, occurring approximately two weeks after the chemotherapy administration, were always considered. If treatment was changed completely, the patients were observed

until the start of a new cancer regimen; if treatment was discontinued, the patients were observed until four weeks after the last cycle, following Extermann et al. [68].

For the toxicity analysis, only severe toxicity (grade ≥ 3) was considered. Overall, hematologic, and nonhematologic toxicity were analyzed separately. The incidence of severe toxicity was investigated for each patient and was assessed (I) for the complete therapy course and (II) at the start of therapy. Both time points were evaluated because severe toxicity is frequently experienced within the first cycle [95]. The "start of therapy" was defined as the first cycle of therapy or, in case of a shorter cycle length, as a minimum of three therapy weeks. Moreover, the time until occurrence of the first severe toxicity was investigated. Because toxicity was documented per cycle, the middle of the cycle was regarded as the time point of toxicity occurrence in this study. Time was documented in weeks.

3.1.6.2 Patient-reported symptom burden

Only focusing on physician-reported toxicity does not reveal the whole picture: Subjective adverse events during cancer therapy are at risk of being under-reported by physicians [96]. It was shown that patients themselves report symptoms earlier and more frequently than the treating physician [97]. Patient-reported outcomes (PRO) can complement physician-based toxicity reporting [97]. Therefore, in this evaluation study, the symptom burden of patients was measured by the Patient-Reported Outcomes version of the Terminology Criteria for Adverse Events (PRO-CTCAE). The PRO-CTCAE items were developed by the US National Cancer Institute as standardized tools for measuring symptoms during cancer therapy from a patient's perspective [94]. The PRO-CTCAE was validated for the German language [98] and consists of 78 symptoms which are assessed regarding different attributes (e.g. frequency, severity) [99]. From this PRO-CTCAE item pool, items can be selected as needed. The PRO-CTCAE items used in this evaluation study were selected based on a general core item set covering relevant toxicity during cancer therapy [100]. Thirteen symptoms were assessed in this study, regarding 21 attributes. The attributes comprised severity, frequency, or interference with daily activities (see Table 3-6; for study materials see Appendix C).

Patients rated the PRO-CTCAE items on a 5-point Likert scale, ranging from zero (not at all; never) to four (very; almost always); see Appendix C. As the National Cancer Institute has not

published a scoring manual for the PRO-CTCAE, it was calculated following the validation of the German PRO-CTCAE by Hagelstein et al. [98]. They computed the score following the scoring manual of the EORTC-QLQ C30, a questionnaire for assessing the quality of life from cancer patients, developed by the European Organisation for Research and Treatment of Cancer (EORTC) [101].

Table 3-6 The PRO-CTCAE items of this evaluation study

Symptom	**Attributes**
Dysphagia	Severity
Dry mouth	Severity
Mucositis	Severity, interference with daily activities
Pain	Frequency, severity, interference with daily activities
Anorexia	Severity, interference with daily activities
Constipation	Severity
Diarrhea	Frequency
Nausea	Severity, frequency
Vomiting	Severity, frequency
Insomnia	Severity, interference with daily activities
Fatigue	Severity, interference with daily activities
Dyspnea	Severity
Peripheral sensory neuropathy	Severity, interference with daily activities

First, for each symptom scale, the different attributes were summarized by calculating the raw score (RS) derived from the mean of the corresponding attributes (Equation 3-1) [101]. In case of missing attributes, the raw score was only calculated if $\geq$ 50% of the values were present.

$$RS = \frac{I_1 + I_2 + \ldots + I_n}{n} \qquad \textit{Equation 3-1}$$

I_1 = Value of item 1

I_2 = Value of item 2

I_n = Value of item n

n = Number of items per scale

Subsequently, the raw score of the respective symptoms was linearized following Equation 3-2 [101]. It was transformed on a scale from 0 to 100% where higher values indicated a higher severity of toxicity. A score of ≥ 75% was defined as severe symptom burden.

$$Score = \left\{\frac{RS}{Range}\right\} \cdot 100 \qquad \textit{Equation 3-2}$$

The recall period of the PRO-CTCAE comprises seven days [94]. In this study, the PRO-CTCAE was measured once per cycle. Preferably, it was performed one to two weeks after the start of a cycle, for enhancing the comparability between regimens. The questionnaires were performed orally in a patient interview; patients not being present in the Johanniter Hospital Bonn were contacted by phone. The PRO-CTCAE was validated for different modes of administration. Nevertheless, regarding oral administration modes, it was only investigated for interactive voice response systems [94, 102]. All PRO-CTCAE items were also captured as physician-reported CTCAE in this study, in order to allow for comparison. Permission for using the PRO-CTCAE was gained from the NCI in a material transfer agreement.

3.1.6.3 Alterations of planned treatment

Toxicity-associated alterations of planned treatment were captured from medical records. Alterations which occurred for other reasons (e.g. progress of cancer) were excluded. The toxicity-associated alterations were classified as “discontinuations” if a patient discontinued the planned treatment early without continuing a systemic cancer treatment. All alterations where patients completely changed the therapy regimen and subsequently continued another systemic cancer therapy were regarded as “changes”. “Dose reductions” were defined as reduction of regimen dosages or omitting one of the drugs from the regimen. Alterations were considered “delays” if the regimen was postponed for several days due to toxicity but was not stopped. Discontinuations and changes were summarized as “major alterations”; delays and reductions were classified as “minor alterations”. The middle of the respective cycle (in weeks) was recorded as the time to occurrence of the first alteration of planned treatment.

3.1.7 Statistical analysis

Data entry and statistical analysis were performed using Microsoft® Excel® 2007 (Microsoft Corporation, Redmond, USA) and IBM® SPSS® Statistics Version 25.0 for Windows (IBM Corporation, Armonk, USA). Figures were generated with GraphPad Prism® Version 6.01 (GraphPad Software, San Diego, USA) and IBM® SPSS® Statistics Version 25.0.

Descriptive statistics were performed for summarizing patient characteristics, risk assessment, and outcome results. A mean with standard deviation (SD) or a median with interquartile range (IQR) was calculated, as appropriate. The frequencies were given as numbers and percentages.

For the inductive statistics in the exploratory analyses, a p-value of < 0.05 was considered as statistically significant. Confidence intervals (CI) of 95% were computed.

3.1.7.1 Relationship between risk assessments

For assessing if the CARG score, the CRASH score, and physicians predict consistent risks, the relationship between the CARG and the CRASH score as well as between the scores and the physicians' judgment were analyzed.

The correlation of the CARG and the CRASH score was tested with the two-sided Spearman's rho (r_s) for ordinal data [103] and illustrated by a scatter plot. The score results were treated as continuous variables. The strength of a correlation increases with higher values of r_s. However, a high correlation does not necessarily mean that two ratings have a high agreement [103]. Therefore, the agreement was assessed by (weighted) kappa and Fisher's exact test (exact chi-square test). Kappa is the gold standard for assessing agreement [104, 105]. It accounts for the possibility of agreement occurring by chance; see Equation 3-3 [105].

$$\kappa = \frac{\rho_o - \rho_e}{1 - \rho_e} \qquad \textit{Equation 3-3}$$

κ = kappa coefficient

ρ_0 = proportion of observed agreement

ρ_e = proportion of agreement by chance

For data with more than two categories, the extension of kappa, the weighted kappa (κ_w), is used [106]. The weighted kappa also considers the distance of disagreement: The disagreement of larger distances (for instance between the categories low and high) is assigned a higher weight than for smaller distances (for instance between the category low and mid-low). The linear weighting scheme according to Agresti was applied [107]. As an essential condition of weighted kappa, the number of categories must be equal for both scores. Therefore, the CRASH score was pooled into three categories, in order to have as many categories as the CARG score. Results of different poolings of the CRASH score were investigated as sensitivity analysis (e.g. pooling mid-low/mid-high vs pooling low/mid-low for analysis). (Weighted) Kappa was interpreted according to Landis and Koch [108], see Table 3-7. Higher values indicate better agreement, whereas negative values imply an agreement worse than chance.

Table 3-7 Kappa statistic and strength of agreement, according to Landis and Koch [108]

Kappa statistic	**Strength of agreement**
< 0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost perfect

The actual agreement (not adjusting for agreement occurring by chance) was investigated by assessing the proportion of the respective CRASH score categories in different CARG score categories. This was illustrated in a stacked bar chart and evaluated by Fisher's exact test. Usually, a chi-square test is applied to analyze if categorical data are statistically significantly associated with each other. However, the chi-square test is not accurate if the expected frequencies are less than five [109]. Since the Fisher's exact test estimates the exact probabilities of chi-square statistics and is also precise at lower sample size [109], this evaluation study with limited sample size used the Fisher's exact test for assessing the statistical significance of the agreement.

3.1.7.2 Predictive performance

The predictive performance for severe toxicity was assessed regarding the CARG and the CRASH score (combined, hematologic, and nonhematologic), as well as physicians' judgment. It was analyzed concerning two time frames: For the toxicity incidence during therapy course and at start of therapy. Moreover, the predictive performances for patient-reported symptoms and for toxicity-associated alterations of the planned treatments were investigated. In order to compare the scores with other commonly used predictors, the predictive performance for severe toxicity regarding the predictors age and ECOG performance status were examined as well. The scores were treated as continuous and physicians' judgment as categorical data, allowing to use the maximal information available.

For judging the predictive performance of scores, in general, two essential aspects should be investigated: calibration and discrimination [110–112]. Calibration assesses if the predicted outcome corresponds with the real outcome; discrimination evaluates if a score can differentiate between patients with a certain outcome versus patients without that outcome [110, 112].

To assess calibration, Fisher's exact test and univariate logistic regression were used. The Fisher's exact test analyzed the association between the proportion of patients with severe toxicity and the predicted risk category of the scores [109, 110]. The proportion of patients with severe toxicity per category was illustrated with bar charts. The categories of the scores were pooled into "low vs high" (CARG: low/mid vs high; CRASH: low/mid-low vs mid-high/high) to ensure an adequate number of patients per group. Furthermore, the proportion of patients with severe toxicity per score value from the development studies of the scores was compared with the proportion of patients with severe toxicity per score value in our study. A logistic regression tests the influence of risk factors on a binary outcome, usually using the Wald statistics [109, 113]. In this evaluation study, we computed a logistic regression with Wald statistics to analyze if the score predictions were significantly associated with the incidence of severe toxicity, severe symptom burden, or toxicity-associated alterations of the planned treatments. The score values were treated as continuous variables, physicians' judgment as categorical variable.

To evaluate discrimination, receiver operating characteristic curves (ROC curves) were calculated [114]. In ROC curves, the x-axis displays 1-specificity (false positive rate) and the y-axis sensitivity (true positive rate) [115]. A larger area under the ROC curve (ROC-AUC) indicates a better discrimination because in this case, the score features a low false positive rate while exhibiting a high true positive rate [115]. A ROC curve close to the line of identity corresponds to a score predicting as well as chance alone, whereas a curve close to the left upper corner indicates an almost perfect test. The ROC-AUC can be interpreted according to Carter [116]: 1 = perfect; $\geq$ 0.9 = excellent; $\geq$ 0.8 = good; $\geq$ 0.7 = fair; $\geq$ 0.5 = poor; 0.5 = no value. Furthermore, ROC curves display the trade-off between sensitivity and specificity along the course of the curve. Each point of the curve represents the respective specificity and sensitivity for a certain cut-off value. Cut-off values divide the score results into different categories, e.g. low or high risk. To detect the best cut-off values for a score, the specific cut-offs with a maximum of sensitivity at a maximum of specificity must be determined. For assessing this optimal trade-off between specificity and sensitivity, the Youden Index was applied [117]:

$$J = maximum\,\{sensitivity + specificity - 1\} \qquad \textit{Equation 3-4}$$

3.1.7.3 Time until occurrence of first toxicity

For cancer patients, not only the total occurrence of toxicity but also the time until occurrence is important. If severe toxicity occurs at a later stage, this allows e.g. for a longer period of treatment before dosage would be reduced due to toxicity. The time until occurrence of the first severe toxicity was analyzed by Kaplan-Meier analysis [118, 119]. The log-rank test was used to compare the risk categories regarding statistically significant differences of time to toxicity [120]. In order to allow for a larger patient cohort in each group, the CARG and CRASH score risk categories were also pooled into “low vs high” (CARG: low/mid vs high; CRASH low/mid-low vs mid-high/high). A univariate proportional hazard model (Cox regression) was assumed to investigate the influence of the score categories on the time to toxicity [121, 122]. In addition, a sensitivity analysis for cycles instead of weeks was carried out.

3.1.7.4 Predictive factors of toxicity

For determining which factors are associated with overall, hematologic, and nonhematologic toxicity, a univariate logistic regression was conducted. All single factors of the CARG and CRASH score were tested, as well as eight further patient-specific factors.

3.1.7.5 Comparison of toxicity and symptom burden

To analyze the agreement between severe patient-reported symptom burden and severe toxicity reported by health care providers, kappa statistics was applied (see 3.1.7.1).

3.1.7.6 Missing Data and study drop-out

Incomplete follow-up data bear the risk of misclassification, leading to either over- or underprediction of the predictive performance. Therefore, only those patients whose medical records could be completely collected for follow-up were considered for the outcome analysis regarding severe toxicity and alterations of planned treatment. If no measurements for certain laboratory parameters had been performed during a cycle, this parameter was classified as missing.

3.2 Medication risk analysis

Older patients commonly exhibit higher risks in their medication. A review of medication is considered as an important aspect of a geriatric assessment. Therefore, this analysis sought to complement the evaluation of the onco-geriatric scores by investigating the risks of polymedication, potentially inadequate medication, and potentially relevant drug-drug interactions in older cancer patients.

3.2.1 Study design

This analysis comprises the medication data captured during the pilot study as well as the evaluation study, i.e. patients of both studies were pooled. Inclusion criteria were similar but differed regarding systemic cancer therapy which was not mandatory in the pilot study. To match eligibility criteria, only patients who were actually treated with a systemic cancer treatment were included from the pilot study. The medication was captured from medical records and analyzed at two time points. First, the medication was investigated at the time of admission to hospital for determining the risks in long-term medication which patients experienced even before start of treatment. Additionally, after start of cancer treatment, the medication risks due to cancer therapy were analyzed regarding antineoplastic agents and supportive care medication. An overview of this analysis is illustrated in Figure 3-2.

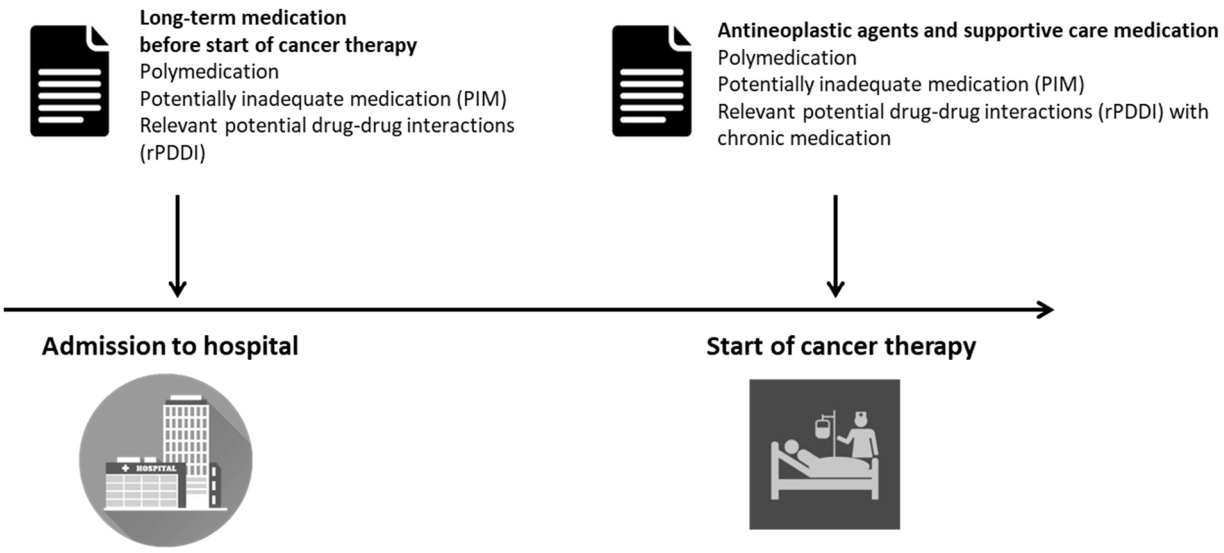


Figure 3-2 Overview of the medication risk analysis

3.2.2 Medication

In general, the medication was counted per active ingredient and not per medicinal product itself. Therefore, if combination medicines comprised for example two active ingredients, they were counted twice. In the following, the term “drug” is always used in terms of “active ingredient”. All active ingredients with systemic effects and an Anatomical Therapeutic Chemical Classification (ATC code) [123] were collected. They were classified according to the

ATC code, level 2 (therapeutic subgroups) [123]. The official ATC index 2017 of the DIMDI (German Institute for Medical Documentation and Information, Deutsches Institut für Medizinische Dokumentation und Information) was used for the ATC classification [123]. If an active ingredient did not have its own ATC code (for example Naloxone with Tilidine), it was not counted separately. Electrolyte solutions (e.g. sodium chloride infusions) or medical gases (e.g. oxygen) were not included. All over-the-counter (OTC) drugs including minerals or vitamins, were considered if they were reimbursable according to the German drug directive (Arzneimittelrichtlinie) [124]. Topical substances (e.g. corticoid ointments), dietary supplements, and medical devices (e.g. sodium chloride nasal sprays) were neglected.

The complete long-term medication which patients took before admission to hospital was included. The medication was recorded from anamnesis in medical records and was only considered if it was continued at least on the first day of admission to hospital. All paused drugs and all drugs just used in case of acute symptoms were excluded because the focus of this analysis was on long-term medication. Risks in long-term medication are usually of higher relevance for patients, compared to drugs just taken once in a while. All anti-infectives (ATC-Code J05) were excluded unless being documented as used for long-term prevention. Different dosages of the same active ingredient were only counted once.

For assessing the antineoplastic agents and supportive care medication, all respective antineoplastic and supportive medication reported on the therapy plan of the first cycle (including rescue medication) was collected and verified in the medical records.

3.2.2.1 Polymedication

In this analysis, polymedication was defined as the concomitant use of ≥ 5 drugs. This cut-off value is commonly used and has shown to be associated with adverse outcome in the elderly [22]. Excessive polymedication (“hyperpolymedication”) was defined as the use of ≥ 10 drugs as discussed by Sharma et al. [22].

3.2.2.2 Potentially inadequate medication

Different criteria can be used to identify potentially inadequate medication (PIM) in older adults. However, none of those criteria has been studied in detail for geriatric cancer patients [125]. A review by Whitman et al. recommends the concomitant use of different screening tools for detecting PIM in older cancer patients [125]: the START/STOPP criteria [35], Beers criteria [31], and MAI criteria [36]. However, implicit tools like the MAI are not reasonable to use in this retrospective setting where comprehensive information about indications or anamnesis was missing [36]. The START/STOPP criteria would require more information as well. The explicit Beers criteria, however, are in particular tailored to the US [31]. Therefore, the EU(7)-PIM list was used in this analysis [33], an explicit PIM list, being objective and appropriate to use in this setting. Moreover, the EU(7)-PIM list also includes recent medication (since being developed in 2015), is widely applicable across Europe and was based on the German PRISCUS list, making it well suited for the application in a German hospital setting. The EU(7)-PIM list states 282 drugs or drug classes from 34 therapeutic groups as PIM [33]. Some drugs are only regarded as PIM under certain conditions, above a certain dose, or duration of treatment. This was also considered for the analysis. However, information was frequently missing for duration. Proton-pump inhibitors (PPI) are regarded as PIM according to the EU(7)-PIM list, if they are taken longer than eight weeks [33]. In this study, all PPI were classified as PIM except if any evidence was found that the PPI was applied for less than eight weeks. For supportive medication during cancer therapy, PIM drugs were only considered if they were applied more than once per cycle for enhancing the clinical relevance of findings.

3.2.2.3 Relevant potential drug-drug interactions

Drug-drug interactions were classified according to the ABDA (Federal Union of German Associations of Pharmacists, Bundesvereinigung Deutscher Apothekerverbände) interaction database [126]. The ABDA interaction database is provided by the Avoxa – Mediengruppe Deutscher Apotheker GmbH [126] and is a commonly used interaction database in German community pharmacies. Since further clinical information was missing, all observed drug-drug interactions were assumed to be potential. For enhancing the clinical impact, this analysis focused on the severe potential drug-drug interactions, the so called “relevant potential drug-drug interactions” (rPDDI). All interactions were defined as relevant which usually require an

intervention or action by health care providers. These included the following five ABDA classifications: “Serious consequences possible – contraindicated”; “serious consequences possible – in certain cases contraindicated”; “serious consequences possible – as precaution contraindicated”; “simultaneous usage not recommended”, or “monitoring/modification needed”. The following three ABDA classifications were excluded: “In some cases monitoring/modification needed”; “monitoring as a precaution”; and “in general no action needed”. The last access to the ABDA interaction database was carried out on 9th and 11th of April 2018 in order to review if any interaction classification had changed.

Regarding cancer therapy, only the rPDDI between the antineoplastic agents or supportive care medication and the long-term medication were included. Interactions between the antineoplastic agents and supportive care medication were excluded because respective cancer therapy regimens are regularly and successfully used in clinical routine. “Desired” rPDDI like methotrexate and folic acid were not counted either. Drugs were only considered for rPDDI if they were applied more than once per cycle for enhancing clinical relevance. If rPDDI occurred with different active ingredients (for example different insulins), the interactions were counted for each substance.

For determining the risk of a drug class being involved as an interaction partner in rPDDI, a prevalence-corrected ratio was calculated. This ratio will be referred to as “interaction propensity”. The interaction propensity was calculated following Equation 3-5:

$$IP = \frac{F_i}{P_d} \qquad \textit{Equation 3-5}$$

IP = Interaction propensity

F_i = Frequency of a drug class being involved as interaction partner in rPDDI

P_d = Prevalence of a drug

3.2.3 Statistical analysis

Descriptive analyses were carried out for polymedication, PIM, and rPDDI regarding long-term medication and antineoplastic agents/supportive care medication. The median and

interquartile range of the number of drugs, PIM, and rPDDI per patient were computed as well as the prevalence of the substances in the patient cohort. For determining whether medication risks in long-term medication were associated with overall, hematologic, and nonhematologic toxicity, a univariate logistic regression was carried out. All types of medication risks in long-term medication (polymedication, PIM, rPDDI) were tested being treated as continuous as well as categorial variables.

4 Results

4.1 Evaluation of onco-geriatric scores

4.1.1 Patient recruitment and follow-up

Patient recruitment was conducted between November 2015 and August 2017 at Johanniter Hospital Bonn; follow-up and data collection were carried out until March 2018. The flow chart of patient recruitment and follow-up is presented in Figure 4-1. In total, 174 patients were assessed for eligibility and 120 (69%) patients were enrolled. Patients mostly refused to participate because they felt the study was too much psychological stress (23/54, 43%) or they experienced physical constraints (13/54, 24%). Six (11%) patients were excluded because of cognitive dysfunction and 6 (11%) patients because they did not receive systemic cancer therapy. One hundred thirteen patients were available for outcome analysis: 3/7 patients could not be evaluated because data from oncology practices were not accessible and 4/7 patients because the site of further therapy was unknown. For all excluded patients, the loss to follow-up occurred early in therapy course: Most patients (5/7) were lost during or after the first cycle; 2/7 patients after the second cycle.

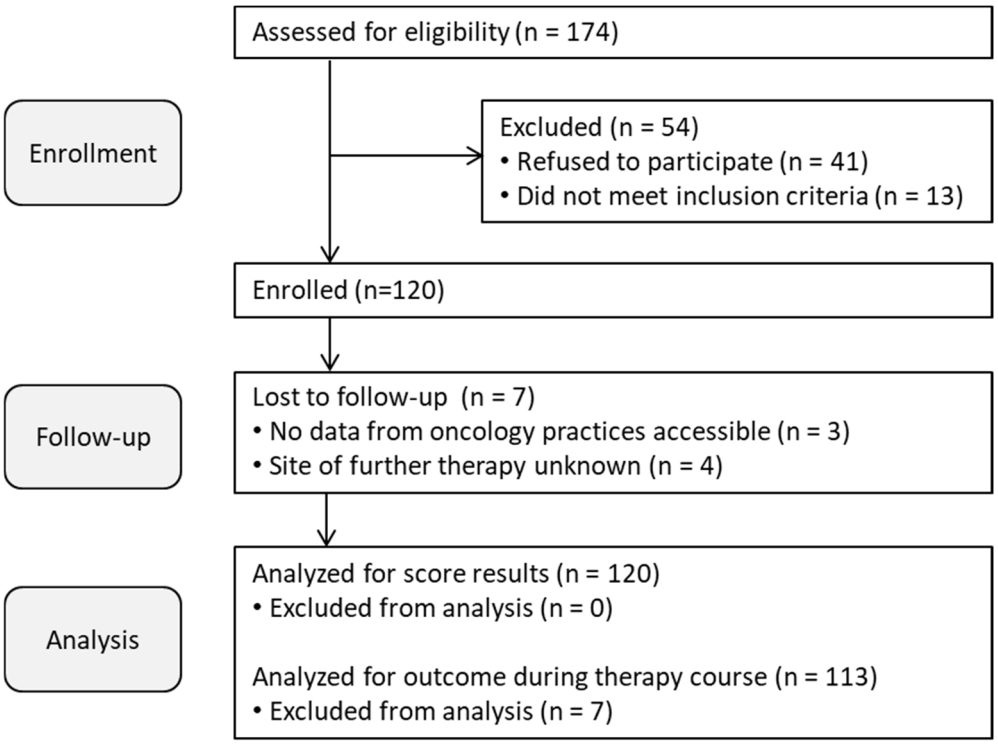


Figure 4-1 Flow chart of patient recruitment

The time between the geriatric assessment and the first therapy cycle was in median 1 day (range 0-53 days). The majority of patients continued with inpatient therapy after the first cycle (63/113, 55.8%), 35/113 (31.0%) with outpatient therapy and 15/113 (13.3%) with in- and outpatient therapy. Most patients were treated at the Johanniter Hospital Bonn during the entire therapy course (87/113, 77.0%). As defined in the study protocol, the follow-up ended after a maximum of six cycles. Forty-seven patients (41.6%) were followed until six therapy cycles. In 66/113 (58.4%) patients, follow-up ended earlier: Twenty-one of sixty-six (31.8%) patients discontinued treatment early (e.g. because of toxicity), 17/66 (25.8%) patients completely changed therapy regimen (e.g. due to progress or toxicity), and 14/66 (21.2%) patients reached the scheduled end of therapy after less than six cycles. In median, 4 cycles were observed during a median of 11 weeks (range 1-45). Eleven patients out of the 113 patients (9.7%) died during follow-up, mostly due to infections or multi-organ failure.

4.1.2 Patient characteristics

The patient characteristics of the study cohort are presented in Table 4-1. Patients had a mean age of 77.2 years at inclusion into the study (SD 4.5, range 70-88) and the cohort was equally distributed between male and female (female 60/120, 50.0%). The study cohort did not represent a typical frail population: Most patients experienced no or little comorbidity according to the CCI (mean 1.1, SD 1.17, range 0-6) and were fully active or at least capable of all self-care according to the ECOG performance status (ECOG 0-2: 105/120, 87.5%). The most frequent CCI conditions comprised diabetes (22/120, 18.3%) and secondary solid tumors (16/120, 13.3%). Renal function was on average mildly decreased (mean 65.6, SD 21.4). Patients took a high number of drugs even before start of cancer therapy (mean 5.1, SD 3.7, range 0-18).

Cancer-related patient characteristics are illustrated in Table 4-2. More than half of patients experienced solid tumors (68/120, 56.7%); the most frequent tumor entities were lung cancer (29/120, 24.2%) and lymphoma (33/120, 27.5%). The majority of patients was treated with chemotherapy (72/120, 60.0%) or a combination of chemotherapy and targeted or immunotherapy (41/120, 34.2%). The most frequent therapy regimens were weekly carboplatin/paclitaxel (16/120, 13.3%) and (R)-CHOP (15/120, 12.5%). Thirty-seven of hundred twenty (30.8%) patients received concomitant radiotherapy.

*Table 4-1 Patient characteristics of the evaluation study at inclusion (n = 120); ECOG, Eastern Cooperative Oncology Group; BMI, body mass index; *for normalized body surface area (BSA) per 1.73 m^2*

	n	%
Age [years]		
70-74	37	30.8
75-79	47	39.2
80-84	26	21.7
≥ 85	10	8.3
Sex		
Female	60	50.0
Male	60	50.0
BMI, WHO		
Underweight (< 18.5 kg/m^2)	4	3.3
Normalweight (18.5-24.9 kg/m^2)	58	48.3
Overweight (25-29.9 kg/m^2)	44	36.7
Obese (≥ 30 kg/m^2)	14	11.7
ECOG performance status		
Fully active (0)	40	33.3
Capable of all self-care (1-2)	65	54.2
Limited or no self-care (3-4)	15	12.5
Charlson Comorbidity Index		
No comorbidity (0)	51	42.5
Little Comorbidity (1-2)	57	47.5
Moderate comorbidity (3-4)	11	9.2
High comorbidity (≥ 5)	1	0.8
Stages of renal insufficiency*		
Normal/high (≥ 90)	10	8.3
Mildly decreased (60-89)	58	48.3
Mildly/moderately decreased (40-59)	26	21.7
Moderately/severely decreased (30-44)	22	18.3
Severely decreased (15-29)	2	1.7
Kidney failure (< 15)	2	1.7

Table 4-1 (continued)

	n	%
Polymedication (long-term medication before start of therapy)		
No polymedication (< 5)	58	48.3
Polymedication (≥ 5-9)	48	40.0
Hyperpolymedication (≥ 10)	14	11.7

*Table 4-2 Cancer-related patient characteristics (n = 120); CUP, cancer of unknown primary; *categorized by body location according to the National Cancer Institute (NCI)*

	n	%
Tumor type		
Solid tumors	68	56.7
Hematological tumors	52	43.3
Tumor entity*		
Respiratory	29	24.2
Lung	29	24.2
Hematological	52	43.3
Lymphoma	33	27.5
Leukemia	11	9.2
Multiple myeloma/plasma cell neoplasm	7	5.8
Myeloproliferative neoplasm	1	0.8
Gynecological	3	2.5
Endometrium	2	1.7
Ovarial	1	0.8
Genitourinary	3	2.5
Urothel	2	1.7
Renal cell	1	0.8
Unknown primary	4	3.3
CUP	4	3.3
Musculoskeletal	1	0.8
Sarcoma	1	0.8

Table 4-2 (continued)

	n	%
Digestive/gastrointestinal	15	12.5
Colorectal	5	4.2
Esophageal	4	3.3
Pancreatic	3	2.5
Gastric	1	0.8
Bile	1	0.8
Breast	11	9.2
Breast	11	9.2
Neuroendocrine	1	0.8
Neuroendocrine	1	0.8
Germ cell	1	0.8
Testicular	1	0.8
Metastasis		
No	24	20.0
Yes	42	35.0
Not applicable/missing	54	45.0
Relapse		
No	104	86.7
Yes	16	13.3
Cancer stage		
I	7	5.8
II	10	8.3
III	29	24.2
IV	58	48.3
Missing	16	13.3
Treatment type		
Chemotherapy	72	60.0
Targeted or immunotherapy	7	5.8
Combined chemotherapy and targeted or immunotherapy	41	34.2

Table 4-2 (continued)

	n	%
Therapy regimen		
Carboplatin/paclitaxel weekly	16	13.3
(Rituximab)-CHOP(cyclophosphamide, doxorubicin, vincristine, prednisone)	15	12.5
Carboplatin/paclitaxel 3-weekly	10	8.3
Bendamustine/rituximab	9	7.5
Mini-(rituximab)-CHOP	6	5.0
Decitabine	5	4.2
Cisplatin/etoposide	4	3.3
Others	55	45.9
Treatment intention		
Palliative	62	51.7
Curative	48	40.0
Others	8	6.7
Missing	2	1.7
Additional therapy		
None	69	57.5
Radiotherapy	31	25.8
Surgery	14	11.7
Radiotherapy and surgery	6	5.0

4.1.3 Risk assessment

4.1.3.1 The CARG score

For the CARG score, a median of 9 (IQR 4, range 4-20) was obtained. Most patients were classified as mid category (61/120, 50.8%) and as high category (52/120, 43.3%). Only 7/120 (5.8%) patients were categorized as low. The CARG score results are presented in Figure 4-2. Results of the CARG score items are illustrated in Table 4-3.

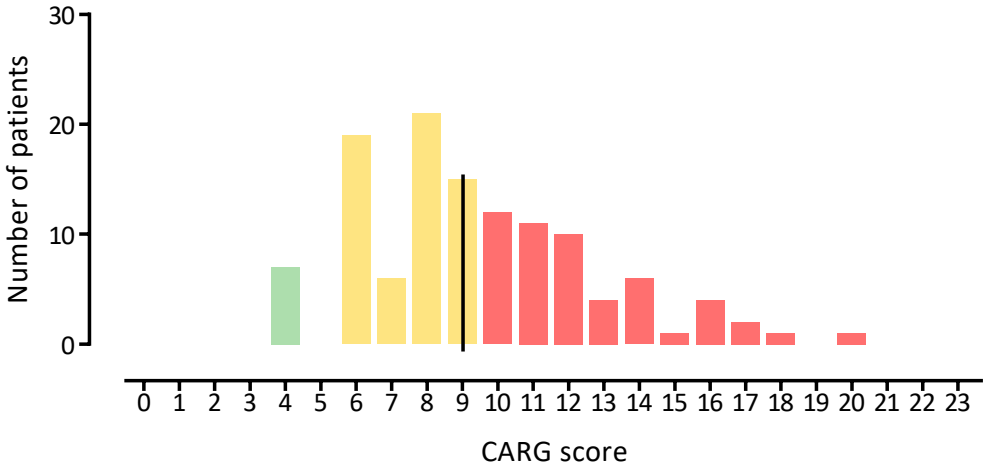


Figure 4-2 *Distribution of the CARG score toxicity predictions; solid line shows median of score results (n = 120); green: low risk; yellow: mid risk; red: high risk*

Table 4-3 *Items of the CARG score (n = 120); GI, gastrointestinal; GU, genitourinary*

	n	%
Socio-demographics		
Age [years]		
≥ 72	111	92.5
< 72	9	7.5
Tumor/treatment variables		
Cancer type		
GI/GU tumor	19	15.8
Other	101	84.2
Dose		
Reduced	17	14.2
Standard	103	85.8
Number of treatment agents		
Monotherapy	18	15.0
Polytherapy	102	85.0
Laboratory variables		
Hemoglobin [g/dL]		
≥ 10 (female), ≥ 11 (male)	79	65.8
< 10 (female), < 11 (male)	41	34.2

Table 4-3 (continued)

	n	%
Creatinine clearance Jeliffe [mL/min] [85]		
< 34	12	10.0
≥ 34	108	90.0
Geriatric assessment variables		
Hearing abilities		
Good/excellent	77	64.2
Fair/worse	43	35.8
Falls in past six months		
0	96	80.0
≥ 1	24	20.0
Medication intake		
No assistance	117	97.5
Requires assistance	3	2.5
Limited in walking one block		
Not limited at all	65	54.2
Limited	55	45.8
Decreased social activity because of health/emotional problems		
A little or none of the time	83	69.2
Some, most, or all of the time	37	30.8

4.1.3.2 The CRASH score

The combined CRASH score exhibited a median of 8 (IQR 2, range 2-11). Patients were mostly stratified as mid-high (72/120, 60.0%). Twenty-two of hundred twenty (18.3%) patients were classified as mid-low and 23/120 (19.2%) as high. Only 3/120 (2.5%) patients were categorized as low. The hematologic CRASH score showed a median of 4 (IQR 2, range 0-6). The majority was classified as mid-high (71/120, 59.2%) and mid-low 40/120 (33.3%). Only 5/120 (4.2%) patients were stratified as low and 4/120 (3.3%) as high. The nonhematologic CRASH score exhibited a median of 6 (IQR 2, range 0-8). Patients were mostly categorized as mid-high (68/120, 56.7%), 24/120 (20.0%) patients as mid-low and 22/120 (18.3%) as high. Only few patients were classified as low (6/120, 5.0%).

The CRASH score results are illustrated in Figure 4-3; the CRASH score items are displayed in Table 4-4.

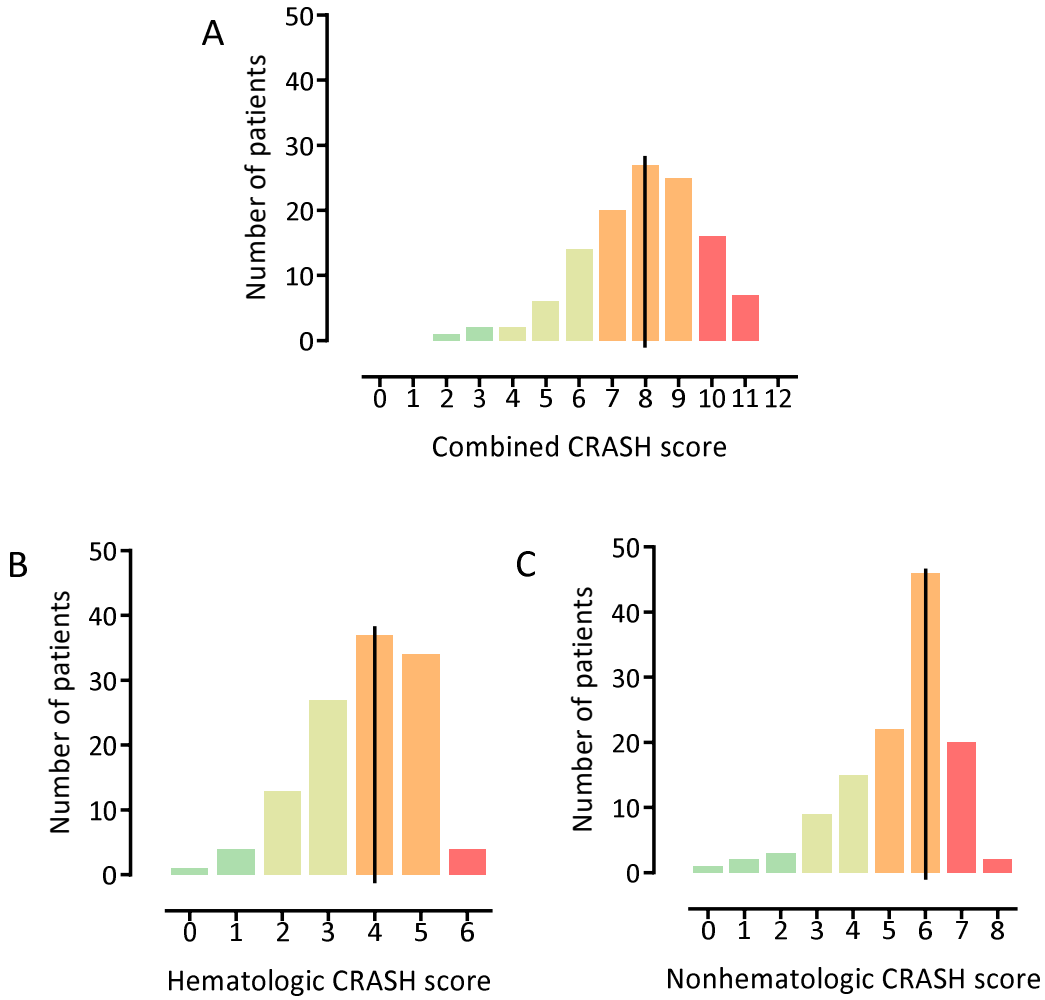


Figure 4-3 *Distribution of the combined (A), hematologic (B), and nonhematologic (C) CRASH score (n = 120); dark green: low risk; light green: mid-low risk; orange: mid-high risk; red: high risk*

Table 4-4 Items of the combined, hematologic, and nonhematologic CRASH score (n = 120); IADL, instrumental activities of daily living; LDH, lactate dehydogenase; ECOG, Eastern Cooperative Oncology Group; MMSE, Mini-Mental State Examination; MNA, Mini Nutritional Assessment; MAX2, chemotherapy toxicity index

	n	%
Hematologic score		
Diastolic blood pressure [mmHg]		
≥ 72	51	42.5
< 72	69	57.5
IADL		
26-29	97	80.8
10-25	23	19.2
LDH [U/L]		
> 167	113	94.2
≤ 167	7	5.8
Nonhematologic score		
ECOG performance status		
0	40	33.3
1-2	65	54.2
3-4	15	12.5
MMSE		
30	22	18.3
< 30	98	81.7
MNA		
28-30	15	12.5
< 28	105	87.5
All subscores		
MAX2		
0 (0-0.44)	27	22.5
1 (0.45-0.57)	48	40.0
2 (< 0.57)	45	37.5

4.1.3.3 Physicians' judgment

Hundred eighteen judgments from physicians were available; two physicians' judgments were missing because the physicians could not be contacted in time before cancer therapy started. Mostly, physicians estimated a mid toxicity risk for overall (65/118, 55.1%), hematologic (65/118, 55.1%), and nonhematologic (61/118, 51.7%) toxicity. A low toxicity risk was expected more often by physicians than a high risk. The distributions of judgments for different toxicity types were similar. Physicians' judgments are illustrated in Figure 4-4.

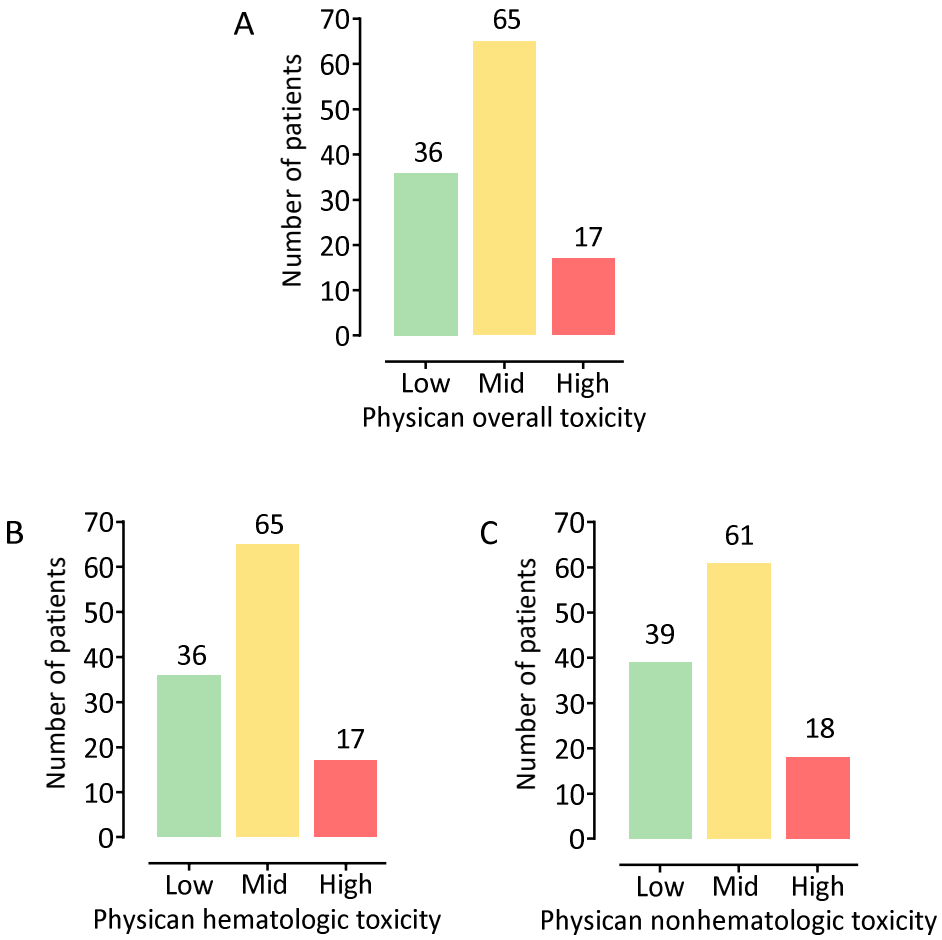


Figure 4-4 *Distribution of physicians' judgments regarding overall (A), hematologic (B), and nonhematologic (C) toxicity risk (n = 118); green: low risk; yellow: mid risk; red: high risk*

4.1.4 Outcome

4.1.4.1 Toxicity

Hundred thirteen patients were available for outcome analysis. Baseline values indicated that 39.8% (45/113) of patients experienced hypertension of grade ≥ 3 toxicity even before start of cancer therapy. For all other types of toxicity, the baseline was generally below grade ≥ 3 toxicity. Hypertension was planned to be assessed since this is a frequent side effect of targeted therapy agents like anti-angiogenic drugs. However, CTCAE criteria definitions are apparently not appropriate for elderly patients who frequently experience hypertension as comorbidity: Due to the strict CTCAE definitions, the baseline prevalence of hypertension grade ≥ 3 was very high. Thus, by including hypertension, the results would have been diluted by "severe" adverse events which are common in the special population of older patients. In order to receive results with a higher clinical relevance, hypertension was neglected for toxicity analysis.

Only a low percentage of laboratory data was missing for follow-up; a median of 2.8% (range 0-19.4%) was not available per patient.

The majority of patients showed signs of overall, hematologic, and nonhematologic grade ≥ 3 toxicity during therapy course. Hematologic toxicity occurred more often than nonhematologic toxicity. The respective findings are displayed in Figure 4-5.

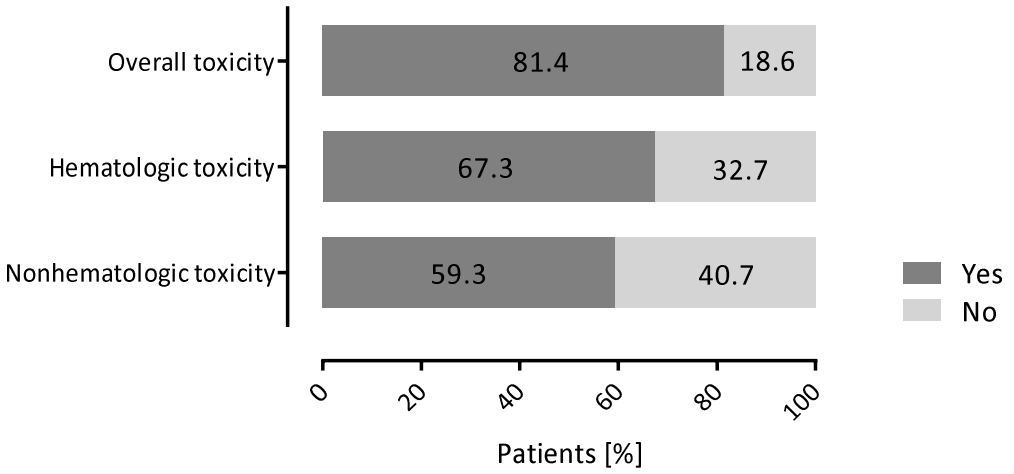


Figure 4-5 Percentage of patients experiencing grade ≥ 3 toxicity during therapy course (n = 113)

The most frequent hematologic grade ≥ 3 toxicity was leukopenia (54/113, 47.8%); the most frequent nonhematologic grade ≥ 3 toxicity comprised infections (37/113, 32.7%). Details regarding the different types of toxicity with corresponding CTCAE grades 3-5 are presented in Table 4-5. Details on the toxicity distribution per patient characteristics and score items are illustrated in Appendix D, Table D-1 to Table D-3.

Table 4-5 Toxicity incidence during therapy course per CTCAE grades 3-5 (n = 113); AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma-glutamyl transferase

	Grade ≥ 3		**Grade 3**		**Grade 4**		**Grade 5**	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Overall toxicity	92	81.4	88	77.9	50	44.2	4	3.5
Hematologic toxicity	76	67.3	70	61.9	41	36.3	0	0.0
Anemia	46	40.7	46	40.7	0	0.0	0	0.0
Febrile Neutropenia	15	13.3	15	13.3	0	0.0	0	0.0
Neutropenia	52	46.0	24	21.2	35	31.0	0	0.0
Leukopenia	54	47.8	33	29.2	32	28.3	0	0.0
Thrombopenia	20	17.7	10	8.8	13	11.5	0	0.0
Nonhematologic toxicity	67	59.3	66	58.4	14	12.4	4	3.5
Acute coronary syndrome	4	3.5	2	1.8	0	0.0	2	1.8
Atrial flutter	1	0.9	1	0.9	0	0.0	0	0.0
Heart failure	3	2.7	2	1.8	1	0.9	0	0.0
Thromboembolic event	3	2.7	3	2.7	0	0.0	0	0.0
AST	10	8.8	9	8.0	1	0.9	0	0.0
ALT	8	7.1	7	6.2	1	0.9	0	0.0
GGT	18	15.9	14	12.4	5	4.4	0	0.0
Bilirubin	1	0.9	1	0.9	0	0.0	0	0.0
Creatinine	5	4.4	4	3.5	2	1.8	0	0.0
Proteinuria	0	0.0	0	0.0	0	0.0	0	0.0
Dyspnea	12	10.6	11	9.7	2	1.8	0	0.0
Erythroderma	0	0.0	0	0.0	0	0.0	0	0.0

Table 4-5 (continued)

	Grade ≥ 3		Grade 3		Grade 4		Grade 5	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Urticaria	0	0.0	0	0.0	0	0.0	0	0.0
Palmar-plantar erythrodysesthesia syndrome	0	0.0	0	0.0	0	0.0	0	0.0
Gastrointestinal bleeding	2	1.8	1	0.9	1	0.9	0	0.0
Infections	37	32.7	32	28.3	3	2.7	2	1.8
Dehydration	6	5.3	6	5.3	0	0.0	0	0.0
Hyponatremia	11	9.7	11	9.7	0	0.0	0	0.0
Hypokalemia	11	9.7	10	8.8	1	0.9	0	0.0
Dysphagia	2	1.8	2	1.8	0	0.0	0	0.0
Dry mouth	0	0.0	0	0.0	0	0.0	0	0.0
Mucositis	2	1.8	2	1.8	0	0.0	0	0.0
Pain	15	13.3	15	13.3	0	0.0	0	0.0
Anorexia	10	8.8	10	8.8	0	0.0	0	0.0
Constipation	0	0.0	0	0.0	0	0.0	0	0.0
Diarrhea	5	4.4	4	3.5	1	0.9	0	0.0
Nausea	6	5.3	6	5.3	0	0.0	0	0.0
Vomiting	5	4.4	4	3.5	1	0.9	0	0.0
Insomnia	2	1.8	2	1.8	0	0.0	0	0.0
Fatigue	14	12.4	14	12.4	0	0.0	0	0.0
Peripheral sensory neuropathy	1	0.9	1	0.9	0	0.0	0	0.0

The median time to first grade $\geq$ 3 toxicity was 2 weeks. Most patients (78/113, 69.0%) exhibited overall grade $\geq$ 3 toxicity already shortly after start of cancer therapy (during the first cycle or at least the first three weeks of therapy). Details regarding the toxicity incidence at start of therapy are presented in Appendix D, Table D-4.

4.1.4.2 Patient-reported symptom burden

From the 113 patients available for toxicity follow-up, 100 patients could be contacted at least once for collecting the symptom burden via PRO-CTCAE. Thirteen patients could not be reached at all, usually due to death or deteriorated general health condition directly after the first cycle. Fifty patients were followed during the complete therapy course; in 38 patients, all cycles except one or two were observed; for the remaining patients, more than three cycles were missing. Reasons for the incomplete follow-up were commonly a bad general health condition, difficulties in contacting the patient, or patient's refusal.

At least one severe symptom (PRO-CTCAE symptom ≥ 75%) was reported by 79% of patients during therapy course. Most frequently, patients reported a severe symptom burden for fatigue, anorexia, and dry mouth. The incidence per symptom scale is presented in Table 4-6.

Table 4-6 Number of patients with severe symptom burden (PRO-CTCAE ≥ 75%) per symptom scale

Symptom scale (attributes)	**Number of patients with severe symptom burden**
Dysphagia (severity)	12
Dry mouth (severity)	34
Mucositis (severity, interference with daily activities)	6
Pain (frequency, severity, interference with daily activities)	20
Anorexia (severity, interference with daily activities)	34
Constipation (severity)	19
Diarrhea (frequency)	18
Nausea (severity, frequency)	17
Vomiting (severity, frequency)	10
Insomnia (severity, interference with daily activities)	14
Fatigue (severity, interference with daily activities)	45
Dyspnea (severity)	8
Peripheral sensory neuropathy (severity, interference with daily activities)	9

4.1.4.3 Agreement of toxicity and symptom burden

For all symptoms except dyspnea, severe symptom burden (≥ 75% PRO-CTCAE) was reported more often by patients than severe toxicity (CTCAE grade ≥ 3) was reported by health care providers. A comparison of the proportion of patients with severe symptom burden and the proportion of patients with severe toxicity is depicted in Figure 4-6. Differences were especially obvious for the symptoms constipation and dry mouth where numerous patients reported severe symptoms whereas health care providers did not report any toxicity at all in the medical records. Severe dysphagia was reported approximately six times more often by patients than by health care providers, severe insomnia seven times and severe peripheral sensory neuropathy about nine times more frequently. Kappa values demonstrated only low agreement between PRO-CTCAE and CTCAE (see Table 4-7). According to Landis and Koch, Kappa values below 0.4 indicate fair, below 0.2 slight, and below 0.0 poor agreement [108]. Insomnia, mucositis, and peripheral sensory neuropathy exhibited poor agreement, the other symptoms slight or fair agreement.

Table 4-7 Kappa values for the agreement between the ≥ 75% PRO-CTCAE symptoms and the grade ≥ 3 CTCAE toxicity; if no result is given: no kappa value computable due to a zero value for CTCAE*

Symptom	Kappa*	P value
Dysphagia	0.260	0.013
Dry mouth	-	
Mucositis	- 0.031	1.000
Pain	0.265	0.013
Anorexia	0.070	0.439
Constipation	-	
Diarrhoe	0.147	0.083
Nausea	0.212	0.033
Emesis	0.214	0.077
Insomnia	- 0.019	1.000
Fatigue	0.113	0.130
Dyspnoe	0.321	0.016
Peripheral sensory neuropathy	- 0.018	1.000

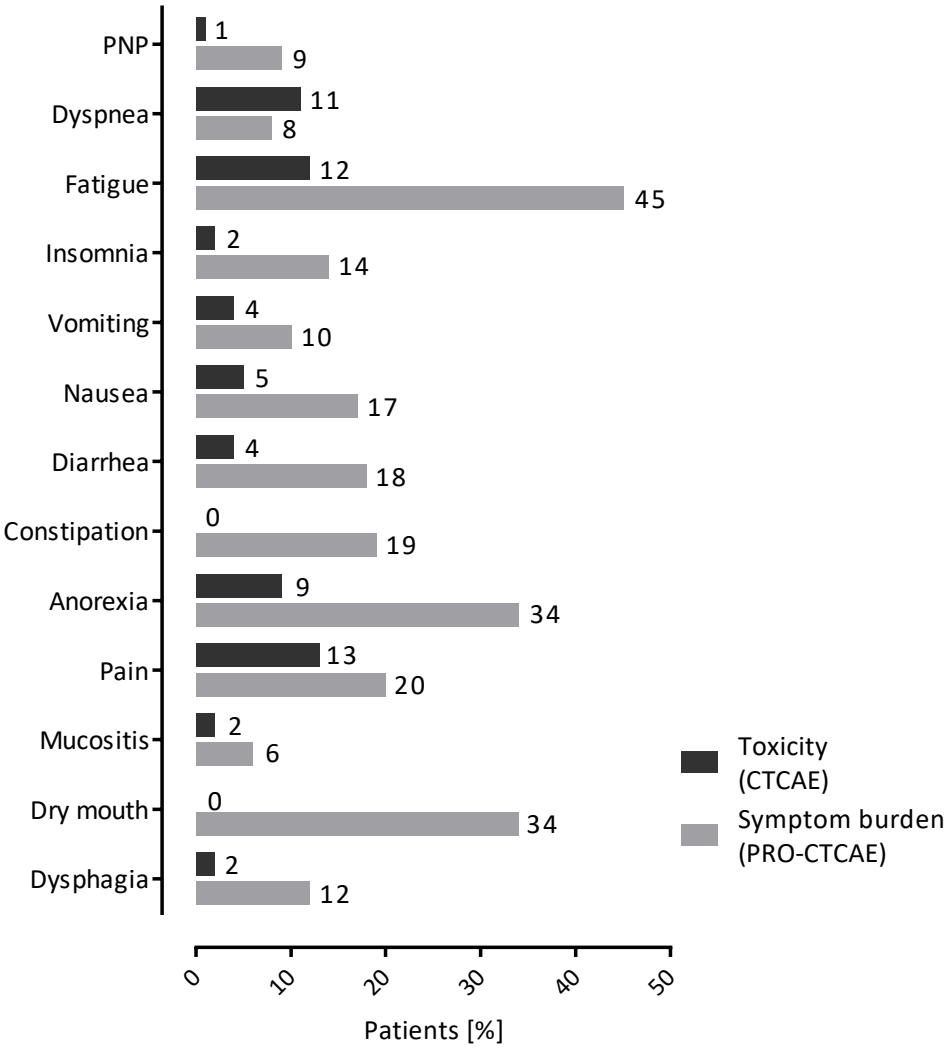


Figure 4-6 Comparison of severe patient-reported symptom burden according to ≥ 75% PRO-CTCAE (n = 100) and severe toxicity reported by health care providers according to ≥ 3 CTCAE (n = 113); PNP, peripheral sensory neuropathy

4.1.4.4 Alterations of planned treatment

Hundred thirteen patients were available for analyzing alterations of the planned treatment. Therapy was discontinued in 31/113 (27.4%) patients after a median of 1 cycle and complete changes of therapy regimen in 11/113 (9.7%) patients after a median of 2 cycles. Therapy delays occurred in 38/113 (33.6%) patients after a median of 1 cycle and therapy reductions in 20/113 (17.7%) patients after a median of 2 cycles. About two-thirds of patients experienced alterations of the planned treatment during therapy course. Major alterations

(changes or discontinuations) occurred in more than one third of patients, minor alterations (delays or dose reductions) in almost half of patients; see Figure 4-7.

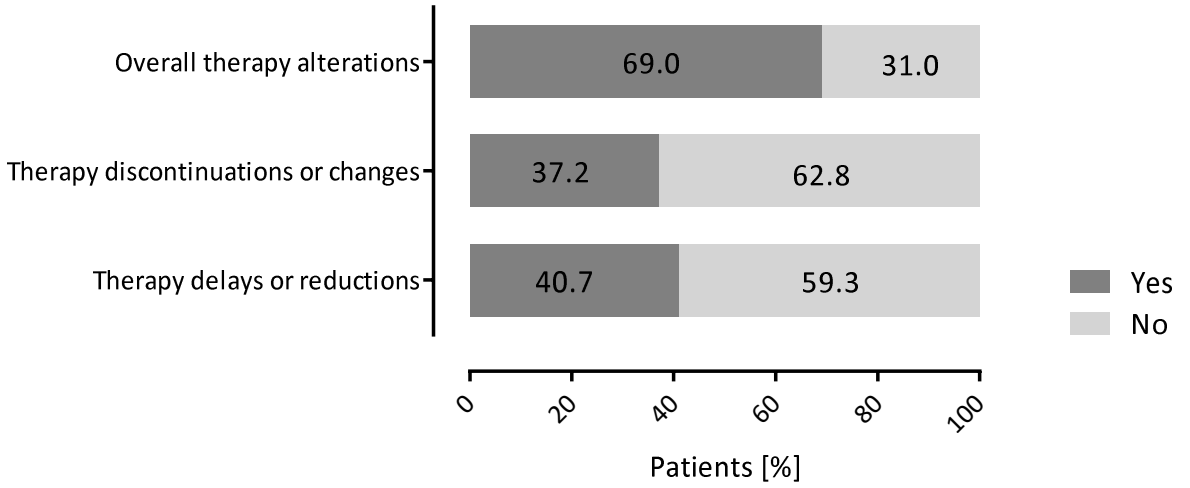


Figure 4-7 Percentage of patients who experienced alterations of their planned treatment during therapy course (n = 113)

Frequent reasons for the alteration of a planned treatment comprised infections, neutropenia/leukopenia, or a deterioration of the general health condition. In 44.2% (50/113) of patients, alterations of the planned treatment already occurred during the first cycle of therapy or during the first three therapy weeks.

4.1.5 Relationship between risk assessments

4.1.5.1 CARG score vs CRASH score

Between the CARG and the combined CRASH score, only a low correlation was found (r_s = 0.203, p = 0.026). Moreover, the chance-corrected agreement between the scores was low, exhibiting a weighted Kappa of 0.057 (CI, -0.074-0.188; p = 0.394). To satisfy the weighted Kappa assumption of equal category numbers, the CRASH score was pooled into three categories by combining the mid-low and mid-high category. Pooling of different categories indicated similar results: The pooled low and mid-low categories exhibited a weighted Kappa of 0.085 (CI, -0.037-0.207; p = 0.141). The proportions of the combined CRASH score categories per CARG score categories did not indicate an association of both scores either

(p = 0.394, Fisher's exact test). These poor results in agreement and correlation indicate that the CARG and the CRASH score predict different risks for patients. The proportions of the combined CRASH score categories per CARG score categories, as well as the correlation of the CARG score with the combined CRASH score are illustrated in Figure 4-8. Correlation and agreement of the hematologic and nonhematologic CRASH score with the CARG score resulted in poor results as well; see Appendix D, Table D-5.

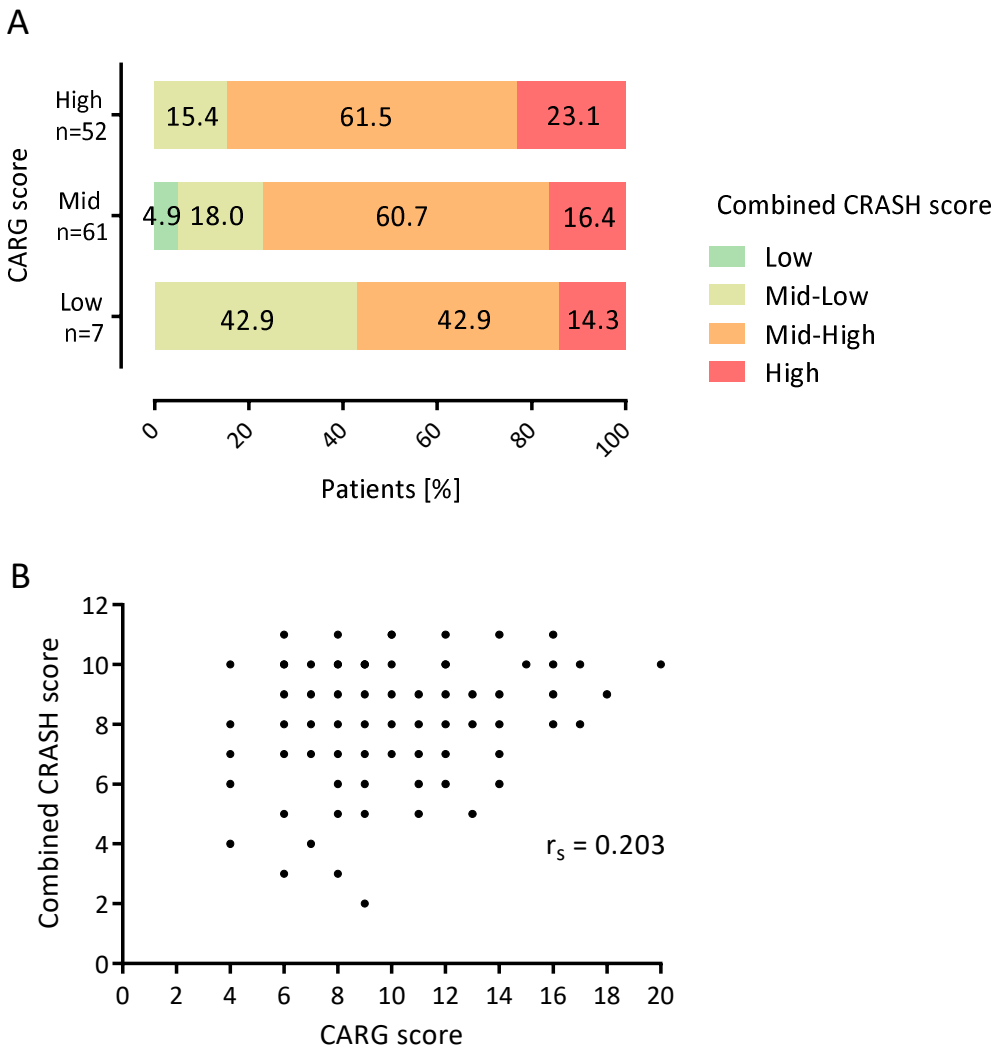


Figure 4-8 Agreement (A) and correlation (B) of the risk categories of the CARG score and the combined CRASH score (n = 120); r_s, Spearman's rho

4.1.5.2 Physicians' judgment vs onco-geriatric scores

The correlation was poor between physicians' judgment for overall toxicity and the CARG score (r_s, 0.011; p = 0.908) as well as between physicians' judgment for overall toxicity and the combined CRASH score (r_s, 0.012; p = 0.897). The chance-corrected agreement yielded poor results as well, both with the CARG score (κ_w, -0.010; CI, -0.115-0.094; p = 0.833) and with the combined CRASH score (κ_w, 0.026; CI, -0.074-0.126; p = 0.589). Fisher's exact test did not suggest a relationship, neither for physicians' judgment with the CARG score (p = 0.133) nor for physicians' judgment with the combined CRASH score (p = 0.734). These poor results in agreement and correlation indicate that physicians and the onco-geriatric scores predict different, complementary risks.

The hematologic and nonhematologic physicians' judgments demonstrated similar poor agreement and correlation for the hematologic and nonhematologic CRASH score, respectively. Details are displayed in Appendix D, Table D-6.

4.1.6 Prediction of severe toxicity by onco-geriatric scores

4.1.6.1 Toxicity during therapy course

Overall toxicity

The CARG and the combined CRASH score exhibited a similar predictive performance for overall grade ≥ 3 toxicity. The risk increased in both scores with higher risk category (CARG: p = 0.051; combined CRASH: p = 0.382; Fisher's exact test). The proportion of patients with grade ≥ 3 toxicity per category is presented in Figure 4-9.

In univariate logistic regression, both scores were found to be significant predictors of toxicity. With each CARG score point, the odds of experiencing overall grade ≥ 3 toxicity increased by about 1.3 (odds ratio (OR), 1.266; CI, 1.048-1.530; p = 0.015); with each CRASH score point, the odds increased by approximately the same factor (OR, 1.337; CI, 1.031-1.734; p = 0.029).

The CARG and the combined CRASH score indicated a similar ROC-AUC of 0.681 (CI, 0.551-0.811; p = 0.010) and 0.650 (CI, 0.519-0.782; p = 0.032), respectively. The ROC curves of both scores are displayed in Figure 4-10.

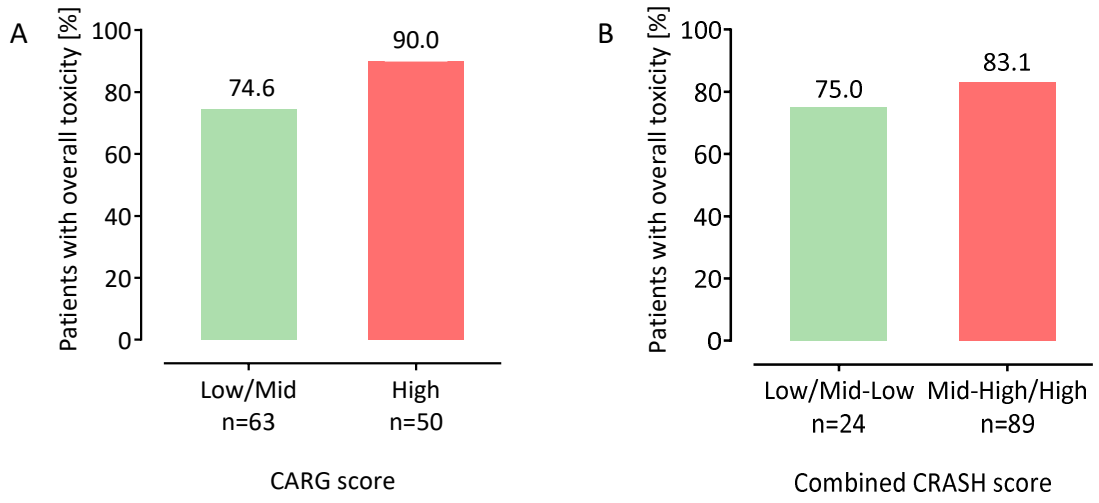


Figure 4-9 *Proportion of patients with grade ≥ 3 toxicity in low vs high categories of the CARG score (A) and the combined CRASH score (B); n = 113*

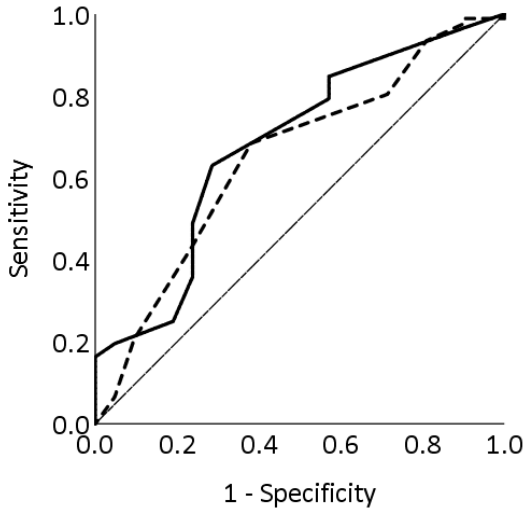


Figure 4-10 *ROC curves of the CARG score and the combined CRASH score for overall grade ≥ 3 toxicity predictions (n = 113); solid line: CARG score; dashed line: combined CRASH score; thin line: line of identity*

The Youden Index determined the optimal cut-off, where a maximum of sensitivity at a maximum of specificity is reached, at ≥ 9 for the CARG score and at ≥ 8 for the combined CRASH score. For the CARG score, at the optimal cut-off, 63.0% of patients with toxicity would be correctly classified as high risk (sensitivity 0.630) and 71.4% of patients without toxicity would be accurately identified as low risk (specificity 0.714). For the combined CRASH score, 68.5% of patients with toxicity would be correctly categorized as high risk (sensitivity 0.685)

and 61.9% of patients without toxicity would be accurately identified as low risk (specificity 0.619).

The proportion of patients with severe toxicity at a certain score value in our study differed from the proportion of patients with severe toxicity at a certain score value in the original development studies of the CARG or CRASH score; see Table 4-8 and Table 4-9. The toxicity risk in our study cohort tends to be higher at a certain score value than being expected based on the percentages of patients with toxicity in the original development cohorts of both scores.

Table 4-8 *Proportion of patients with overall severe toxicity during therapy course per CARG score value in our study compared with the proportion of patients with toxicity per score value in the original CARG development study; * derived from the development study by the CARG online calculator (http://www.mycarg.org)*

CARG score value	Number of patients per score value in our study	Proportion of patients with toxicity in our study [%]	Proportion of patients with toxicity in development study [%]*
4	5	60.0	30
6	18	61.1	44
7	5	100.0	51
8	21	71.4	59
9	14	92.9	66
10	12	100.0	72
11	11	90.9	78
12	8	62.5	82
13	4	75.0	86
14	6	100.0	90
15	1	100.0	92
16	4	100.0	94
17	2	100.0	95
18	1	100.0	97
20	1	100.0	98

*Table 4-9 Proportion of patients with overall severe toxicity during therapy course per combined CRASH score value in our study compared with the proportion of patients with toxicity in the original CRASH development study; * from the CRASH development study* [68]

CRASH score value	Number of patients per score value in our study	Proportion of patients with toxicity in our study [%]	Proportion of patients with toxicity in development study [%]*
2	1	100.0	50
3	2	0.0	50
4	1	100.0	58
5	6	66.6	58
6	14	85.7	58
7	18	61.1	77
8	26	88.5	77
9	24	87.5	77
10	14	92.9	79
11	7	85.7	79

Hematologic toxicity

The hematologic CRASH score indicated a better predictive performance than the CARG score for hematologic grade ≥ 3 toxicity. Using the hematologic CRASH score, toxicity increased with higher risk category (p = 0.002; Fisher's exact test). However, using the CARG score, toxicity increased only slightly (p = 0.687; Fisher's exact test). The proportions of patients with hematologic toxicity per category are presented in Figure 4-11. In univariate logistic regression, only the hematologic CRASH score predicted toxicity significantly (hematologic CRASH: OR, 1.602; CI, 1.135-2.261; p = 0.007; CARG: OR, 1.048; CI, 0.925-1.186; p = 0.462). The hematologic CRASH score exhibited a better ROC-AUC than the CARG score: 0.665 (CI, 0.554-0.776; p = 0.005) and 0.564 (CI, 0.445-0.683; p = 0.271), respectively (see Figure 4-12). The Youden Index determined the optimal cut-off for the CARG score at ≥ 9 and for the hematologic CRASH score at ≥ 4. At this optimal cut-off, for the CARG score, sensitivity was 0.632 and specificity 0.568; for the hematologic CRASH score sensitivity was 0.724 and specificity 0.595. In our study cohort, the percentages of patients with hematologic toxicity per hematologic CRASH score value tended to be higher than the percentages being observed

in the original CRASH development study; see Appendix D, Table D-7. Since the CARG score was not developed for the prediction of hematologic toxicity, no exact percentages of hematologic toxicity were available from the development study for comparison.

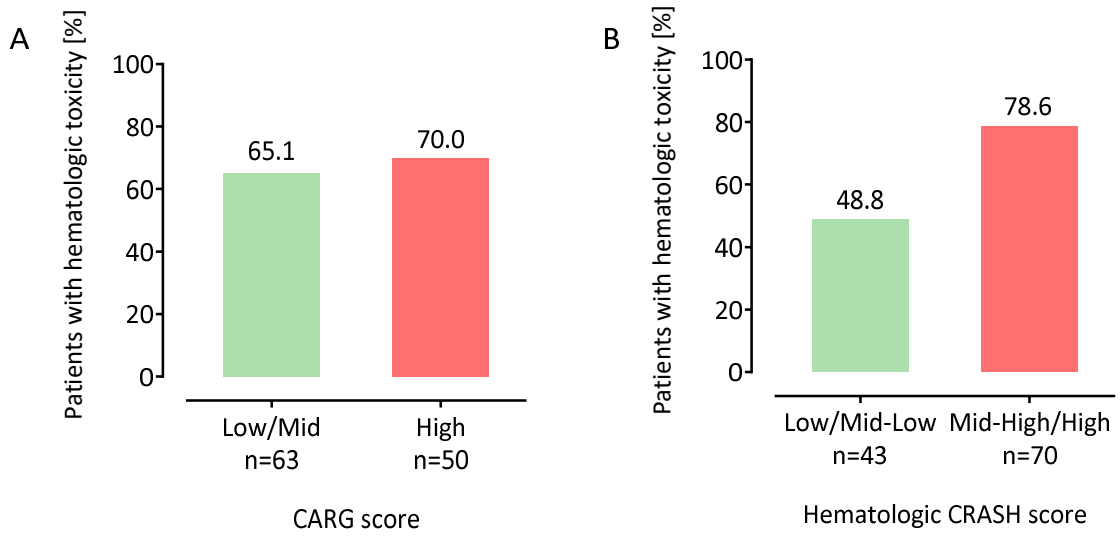


Figure 4-11 Proportion of patients with hematologic grade ≥ 3 toxicity in low vs high categories of the CARG score (A) and the hematologic CRASH score (B); n = 113

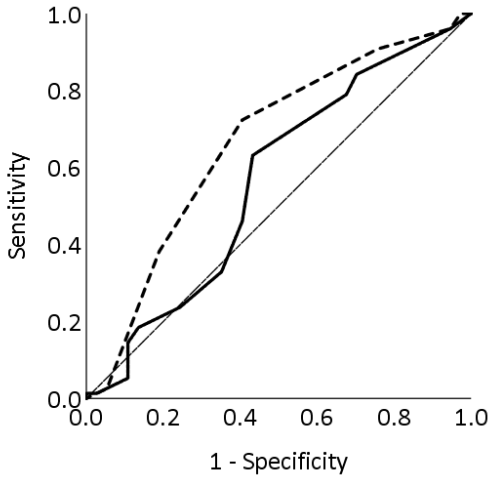


Figure 4-12 ROC curves of the CARG score and the hematologic CRASH score for predictions of hematologic grade ≥ 3 toxicity (n = 113); solid line: CARG score; dashed line: hematologic CRASH score; thin line: line of identity

Nonhematologic toxicity

The CARG and the nonhematologic CRASH score demonstrated a similar predictive performance for nonhematologic grade ≥ 3 toxicity. In both scores, toxicity risk increased with higher risk category (CARG: p = 0.007; nonhematologic CRASH: p = 0.081; Fisher's exact test; Figure 4-13). In univariate logistic regression, both scores were significant predictors of toxicity. The CARG score indicated an OR of 1.219 (CI, 1.062-1.398; p = 0.005), the nonhematologic CRASH score an OR of 1.363 (CI, 1.044-1.781; p = 0.023). The CARG score and the nonhematologic CRASH score denoted a similar ROC-AUC; 0.662 (CI, 0.561-0.763; p = 0.003) and 0.651 (CI, 0.550-0.752; p = 0.007), respectively (Figure 4-14). The Youden Index determined the optimal cut-off for the CARG score at ≥ 10 (sensitivity: 0.552; specificity: 0.717) and for the nonhematologic CRASH score at ≥ 6 (sensitivity: 0.672; specificity: 0.565). A comparison of the nonhematologic toxicity proportions per score value in our study versus the proportions per respective score value in the original CRASH development study are displayed in Appendix D, Table D-8. Since the CARG score was not developed for the prediction of nonhematologic toxicity, no exact percentages of nonhematologic toxicity were available from the development study for comparison.

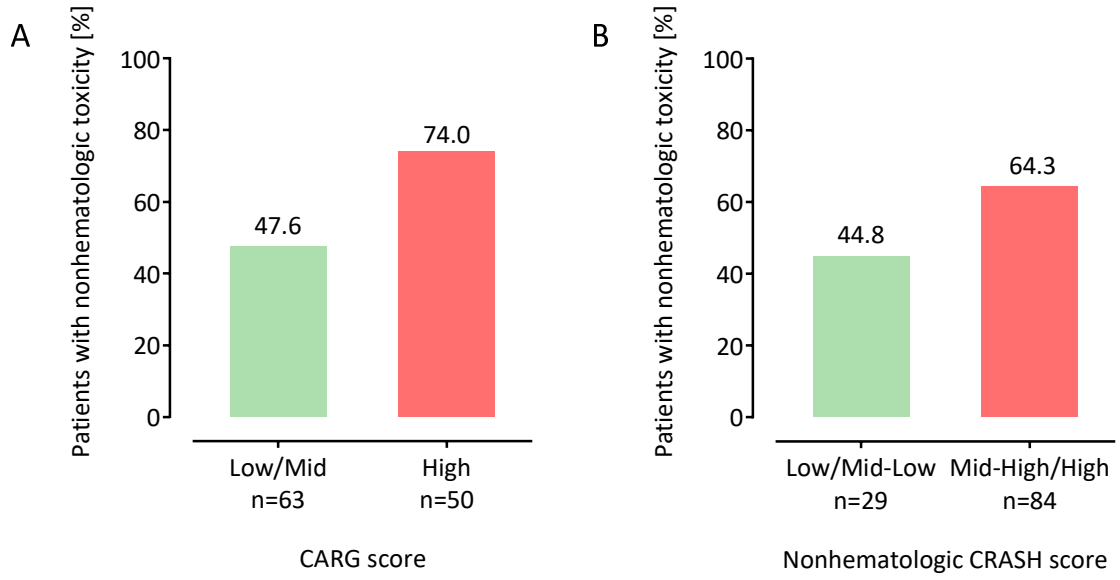


Figure 4-13 Proportion of patients with nonhematologic toxicity grade ≥ 3 in low vs high categories of the CARG score (A) and the nonhematologic CRASH score (B); n = 113

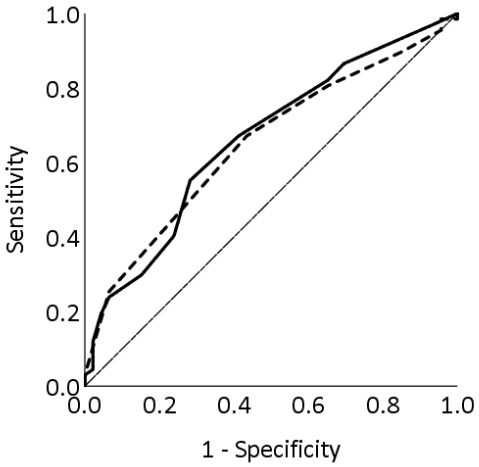


Figure 4-14 ROC curves of the CARG score and the nonhematologic CRASH score for nonhematologic grade ≥ 3 toxicity predictions (n = 113); solid line: CARG score; dashed line: nonhematologic CRASH score; thin line: line of identity

4.1.6.2 Toxicity at start of therapy

Overall toxicity

In line with the results for the complete therapy course, both scores indicated similar predictive performance regarding the prediction of overall grade ≥ 3 toxicity at start of therapy. The risk increased in both scores significantly with higher category (CARG: p = 0.002; combined CRASH: p = 0.044; Fisher's exact test); see Figure 4-15. In univariate logistic regression, the CARG and combined CRASH score both significantly predicted toxicity at start of therapy. The CARG score exhibited an OR of 1.224 (CI, 1.054-1.421; p = 0.008), the combined CRASH score an OR of 1.372 (CI, 1.085-1.735; p = 0.008). Both scores resulted in similar ROC-AUC values (CARG: 0.670, CI 0.562-0.778, p = 0.004; combined CRASH: 0.668, CI 0.559-0.777, p = 0.004); see Figure 4-16.

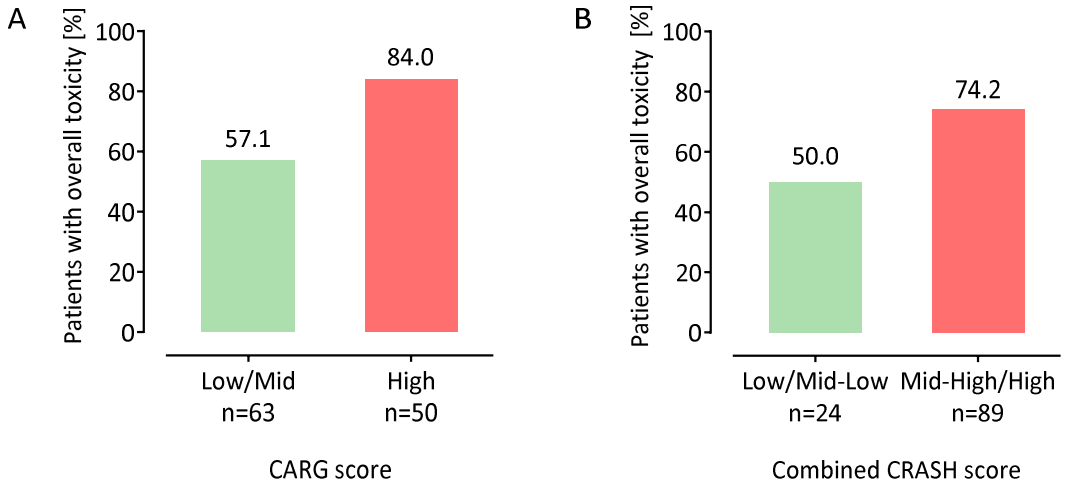


Figure 4-15 Proportion of patients with overall grade ≥ 3 toxicity at start of therapy in low vs high categories of the CARG (A) and the combined CRASH score (B); n = 113

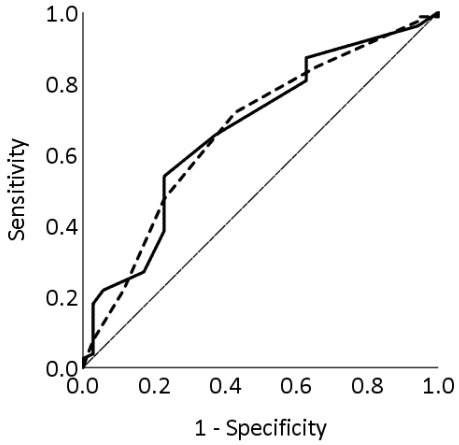


Figure 4-16 ROC curves of the CARG score and the combined CRASH score for overall grade ≥ 3 toxicity predictions at start of therapy (n = 113); solid line: CARG score; dashed line: combined CRASH score; thin line: line of identity

Hematologic toxicity

Contrary to the results for the complete therapy course, predictive performance of the hematologic CRASH was not superior to the CARG score regarding grade ≥ 3 hematologic toxicity at start of therapy. In both scores, the proportion of patients with toxicity at start of therapy increased with higher category. Differences in categories were statistically significant

for the CARG score as well as for the hematologic CRASH score (CARG: p = 0.014; hematologic CRASH: p = 0.034; Fisher's exact test); see Figure 4-17. The univariate logistic regression indicated that neither the CARG score nor the hematologic CRASH score significantly predicted hematologic toxicity at start of therapy (CARG: OR 1.116, CI 0.991-1.258, p = 0.071; hematologic CRASH: OR 1.300, CI 0.948-1.783, p = 0.103). The ROC-AUC of the hematologic CRASH score was slightly lower than the ROC-AUC of the CARG score (hematologic CRASH: 0.592, CI 0.486-0.698, p = 0.092; CARG: 0.638, CI 0.534-0.742, p = 0.012); see Figure 4-18.

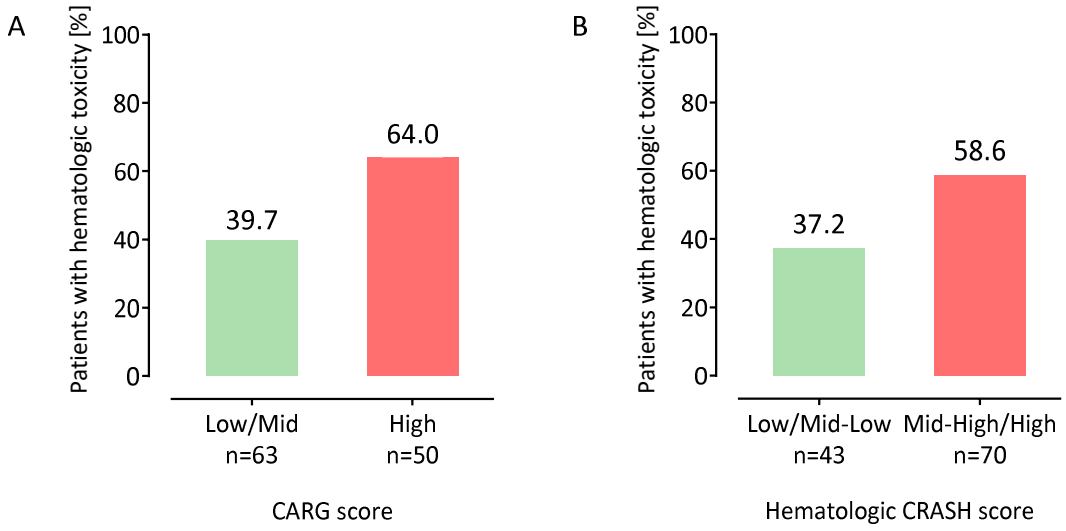


Figure 4-17 Proportion of patients with hematologic grade ≥ 3 toxicity at start of therapy in low vs high categories of the CARG score (A) and the hematologic CRASH score (B); n = 113

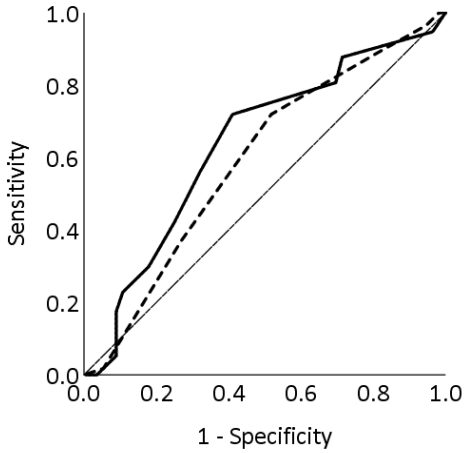


Figure 4-18 ROC curves of the CARG score and the hematologic CRASH score for hematologic grade ≥ 3 toxicity predictions at start of therapy (n = 113); solid line: CARG score, dashed line: hematologic CRASH score; thin line: line of identity

Nonhematologic toxicity

The results for the grade ≥ 3 nonhematologic toxicity at start of therapy resembled the findings obtained for the complete therapy course. Incidence of nonhematologic toxicity increased in both scores with higher category (CARG: p = 0.023; nonhematologic CRASH: p = 0.084; Fisher's exact test); see Figure 4-19. Both scores exhibited significant predictions in univariate logistic regression (CARG: OR 1.162, CI 1.027-1.315, p = 0.018; nonhematologic CRASH: OR 1.351, CI 1.029-1.774, p = 0.030). Both ROC-AUC were similar as well: The CARG score indicated a ROC-AUC of 0.629 (CI, 0.527-0.732; p = 0.018) and the nonhematologic CRASH score a ROC-AUC of 0.633 (CI, 0.531-0.735; p = 0.015); see Figure 4-20.

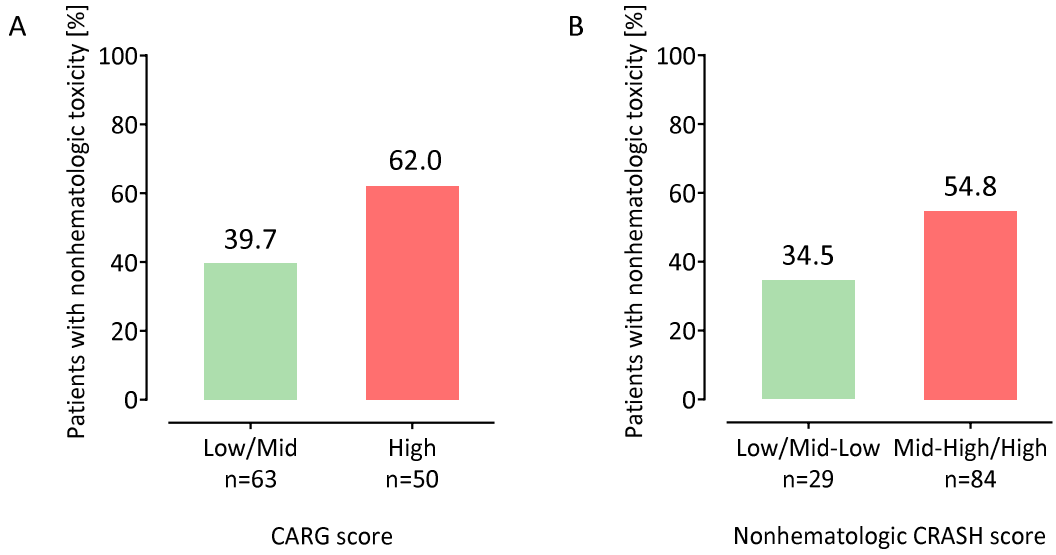


Figure 4-19 Proportion of patients with nonhematologic grade ≥ 3 toxicity at start of therapy in low vs high categories of the CARG score (A) and the nonhematologic CRASH score (B); n = 113

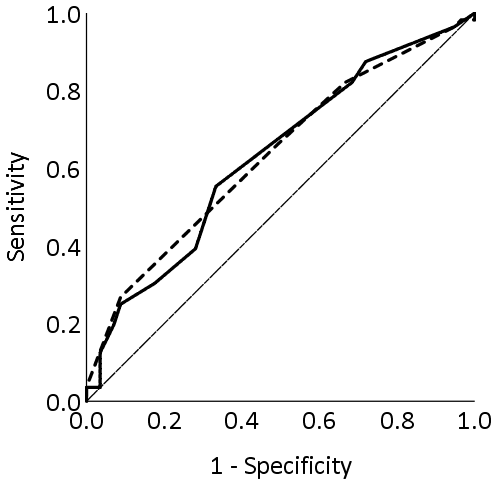


Figure 4-20 ROC curves of the CARG score and the nonhematologic CRASH score for nonhematologic grade ≥ 3 toxicity predictions at start of therapy (n = 113); solid line: CARG score, dashed line: nonhematologic CRASH score; thin line: line of identity

4.1.6.3 Time-related prediction

Overall toxicity

A difference in time until onset of overall grade ≥ 3 toxicity was found between low and high categories of the CARG score (low/mid vs high) as well as the combined CRASH score (low/mid-low vs mid-high/high). In the high category, toxicity occurred faster than in the low category (see Figure 4-21 and Figure 4-22). This finding was statistically significant for the CARG score (p = 0.001, log-rank test), however, not for the combined CRASH score (p = 0.063, log-rank test). The Kaplan-Meier curves of the CRASH score crossed after approximately 20 weeks of therapy but this was not further considered due to the very low patient numbers in each category at this point. In total, half of the patients experienced severe toxicity after 2 weeks of therapy (CI, 1.701-2.299). For the CARG score, the median time to toxicity resulted in 3 weeks (CI, 2.243-3.757) for the low category and 1.75 weeks (CI, 1.483-2.017) for the high category. The risk of experiencing severe toxicity almost doubled in the high CARG score category compared to the low category: Cox regression resulted in a hazard ratio (HR) of 1.908 (CI, 1.250-2.911, p = 0.003). The combined CRASH score indicated a median time to toxicity of 4 weeks (CI, 0.000-8.019) for the low category and 2 weeks (CI, 1.858-2.142) for the high category. The high category exhibited a 1.6-fold risk of toxicity (HR, 1.590; CI, 0.945-2.678;

p = 0.081). The Kaplan-Meier estimates are listed in Appendix D, Table D-9 to Table D-12. An exploratory analysis with cycles as time unit instead of weeks yielded similar results.

Hematologic toxicity

The CARG score as well as the hematologic CRASH score exhibited both a significant difference in the time until hematologic grade ≥ 3 toxicity occurred (CARG: p = 0.019, hematologic CRASH: p = 0.014; log-rank test). Kaplan-Meier plots are illustrated in Figure 4-23 and Figure 4-24. The crossing of the Kaplan-Meier curves in the hematologic CRASH score was neglected due to the very low patient numbers in both categories at this time-point. The median time until onset of hematologic toxicity was in total 3.5 weeks (CI, 2.165-4.835). For the CARG score, half of patients had experienced hematologic toxicity after 5 weeks (CI, 1.839-8.161) in the low category and after 2 weeks (CI, 1.637-2.363) in the high category. For the hematologic CRASH score, half of patients had shown signs of hematologic toxicity after 8 weeks (CI, 0.013-15.987) in the low category and after 3 weeks (CI, 1.844-4.156) in the high category. The hazard ratio of both scores was similar, indicating 1.691 (CI, 1.063-2.692; p = 0.027) for the CARG score and 1.822 (CI, 1.099-3.022; p = 0.020) for the hematologic CRASH score.

Nonhematologic toxicity

In both scores, the nonhematologic grade ≥ 3 toxicity occurred earlier in the high category than in the low category. However, this was only significant for the CARG score (CARG: p = 0.003; nonhematologic CRASH: p = 0.096; log-rank test). The respective Kaplan-Meier plots are presented in Figure 4-25 and Figure 4-26. In general, half of the patients had experienced the first severe nonhematologic toxicity after 4 weeks (CI, 0.453-7.547). For the CARG score, the median time to toxicity was 14 weeks (CI, 1.123-26.877) in the low category and 2 weeks (CI, 1.146-2.854) in the high category. For the nonhematologic CRASH score, the median time until onset of toxicity was 18 weeks (CI, 0.906-35.094) in the low category and 3 weeks (CI, 1.533-4.467) in the high category. Patients in the high CARG category exhibited an almost doubled risk of experiencing nonhematologic toxicity compared to the low category (HR, 1.991; CI, 1.227-3.229; p = 0.005). The hazard ratio regarding the nonhematologic CRASH score categories was slightly lower (HR, 1.638; CI, 0.891-3.009; p = 0.112).

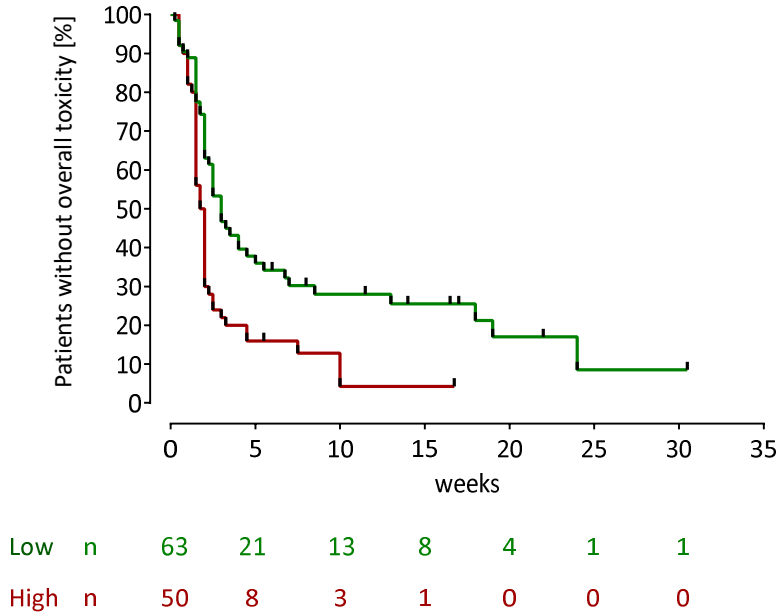


Figure 4-21 *Kaplan-Meier analysis of the time to the first occurrence of severe overall toxicity dependent on the category of the CARG score (low-mid versus high); censored: low 25.4 %, high 10.0%; green line: low CARG score; red line: high CARG score; vertical line: censored data; n = 113*

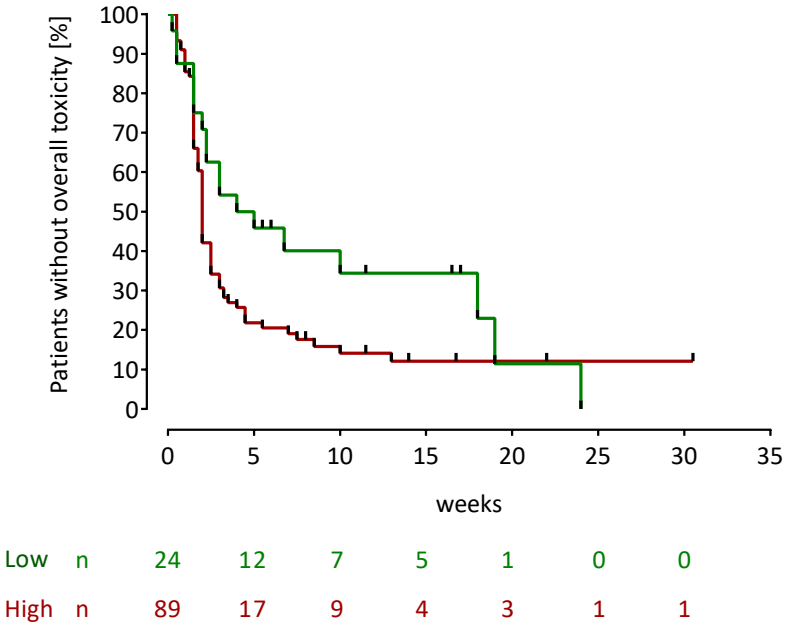


Figure 4-22 *Kaplan-Meier analysis of the time to the first occurrence of severe overall toxicity dependent on the category of the combined CRASH score (low/mid-low versus mid-high/high); censored: low 25.0%, high 16.9%; green line: low CRASH score; red line: high CRASH score; vertical line: censored data; n = 113*

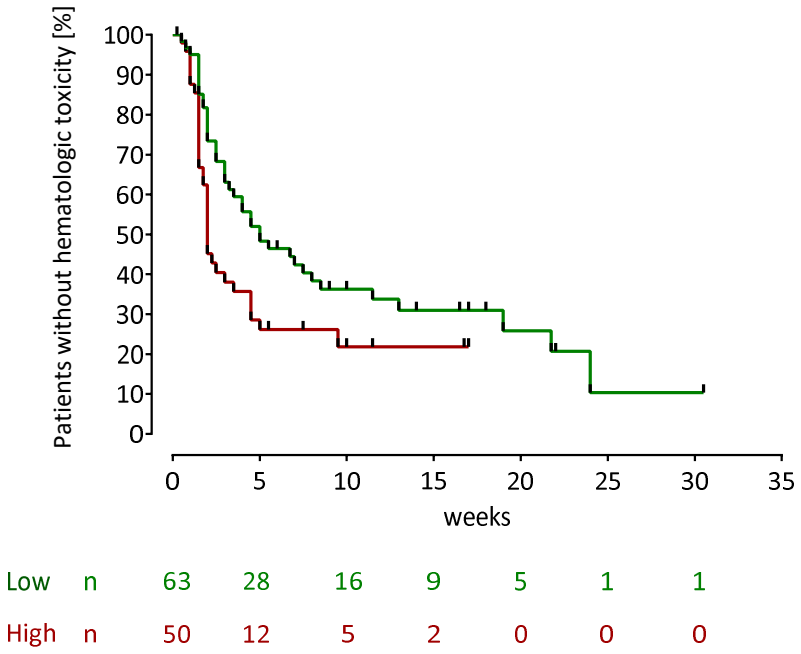


Figure 4-23 *Kaplan-Meier analysis of the time to the first occurrence of severe hematologic toxicity dependent on the category of the CARG score (low-mid versus high); censored: low 34.9%, high 30.0%; green line: low CARG score; red line: high CARG score; vertical line: censored data; n = 113*

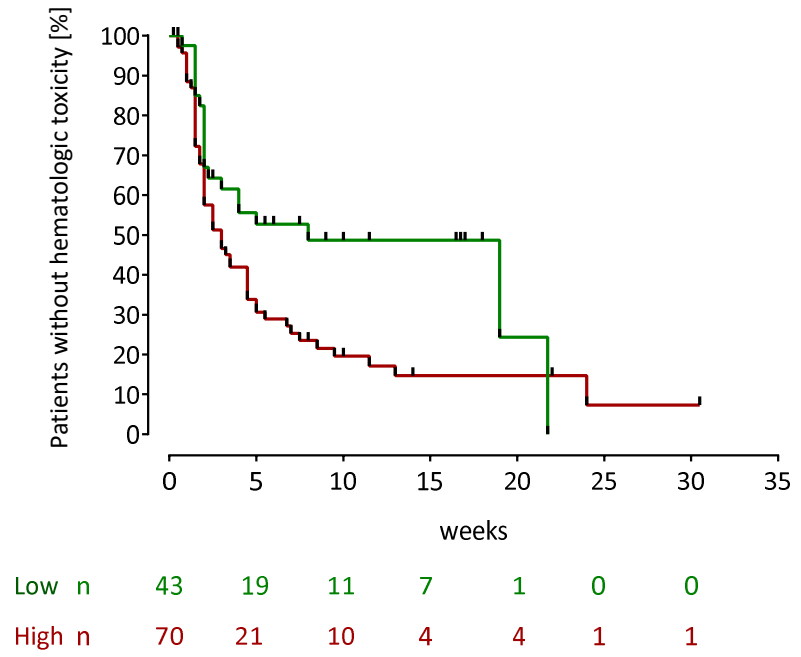


Figure 4-24 *Kaplan-Meier analysis of the time to the first occurrence of severe hematologic toxicity dependent on the category of the hematologic CRASH score (low/mid-low versus mid-high/high); censored: low 51.2%, high 21.4%; green line: low CRASH score; red line: high CRASH score; vertical line: censored data; n = 113*

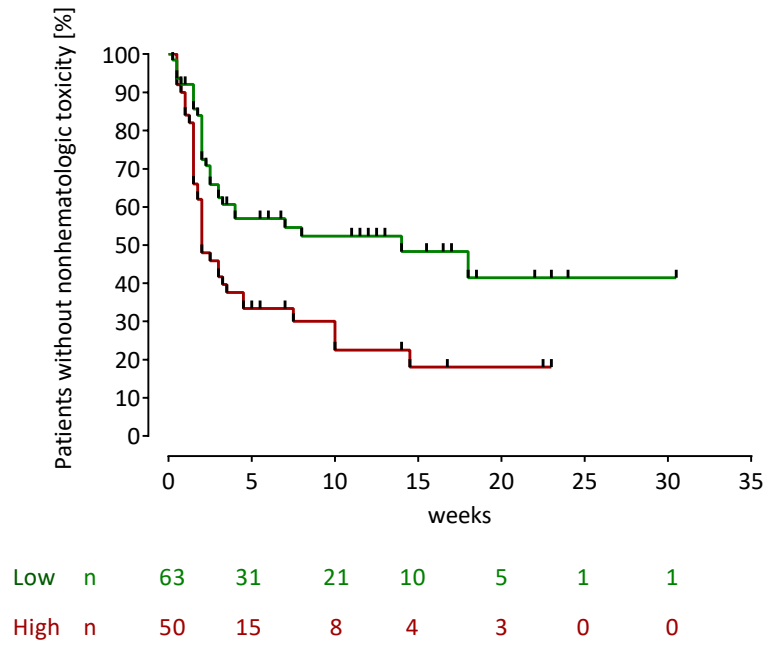


Figure 4-25 *Kaplan-Meier analysis of the time to the first occurrence of severe nonhematologic toxicity dependent on the category of the CARG score (low-mid versus high); censored: low 52.4%, high 26.0%; green line: low CARG score; red line: high CARG score; vertical line: censored data; n = 113*

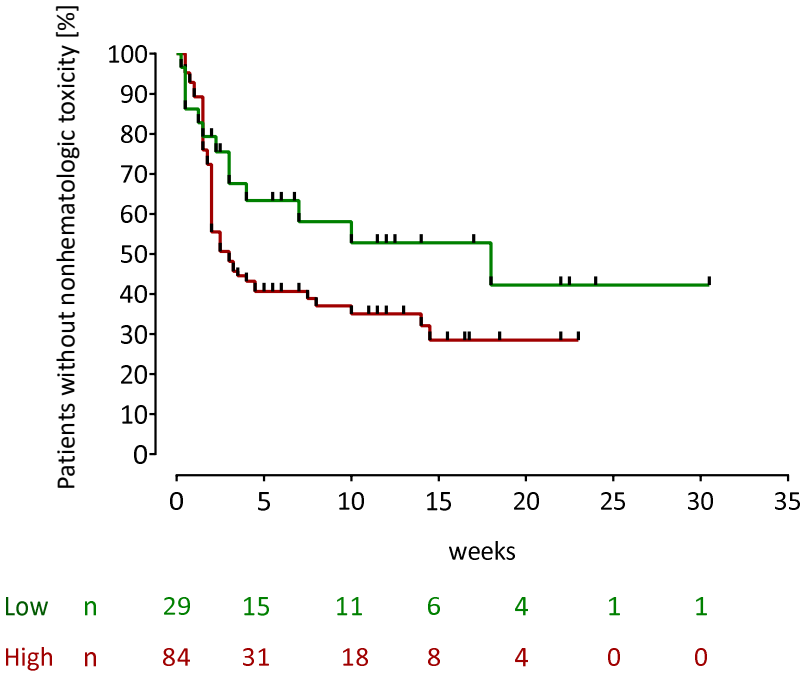


Figure 4-26 *Kaplan-Meier analysis of the time to the first occurrence of severe nonhematologic toxicity dependent on the category of the nonhematologic CRASH score (low/mid-low versus mid-high/high); censored: low, 55.2% and high, 35.7%; green line: low CRASH score; red line: high CRASH score; vertical line = censored data; n = 113*

4.1.7 Prediction of severe toxicity by other predictive factors

4.1.7.1 Physicians' judgment

One hundred eleven patients were available for assessing the predictive performance of physicians' judgment. For two patients, physicians' judgment was missing.

Overall toxicity

The physicians' judgment did not exhibit adequate predictive performance for overall grade ≥ 3 toxicity and resulted in worse predictive performance than the onco-geriatric scores. The toxicity incidence increased only slightly with a higher toxicity risk according to physicians' judgment (p = 0.576, Fisher's exact test); see Figure 4-27. The physicians' judgment did not significantly predict overall toxicity in logistic regression (low vs mid: OR 1.664, CI 0.585-4.731, p = 0.339; low vs high: OR 2.240, CI 0.417-12.042, p = 0.347). The ROC-AUC demonstrated only low discrimination (ROC-AUC, 0.573; CI, 0.433-0.712; p = 0.311); see Figure 4-27. These data indicate that the CARG and the CRASH score are complementary to clinical judgment alone.

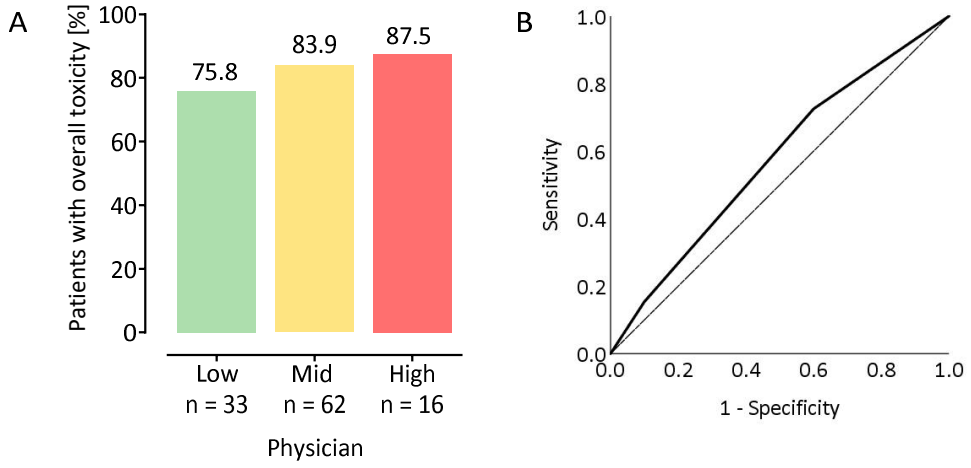


Figure 4-27 (A) Proportion of patients with overall grade ≥ 3 toxicity per physicians' judgment of overall toxicity during therapy course; (B) ROC curve of the physicians' judgment for overall grade ≥ 3 toxicity during therapy course; solid line: physicians' judgment; thin line: line of identity; n = 111

Hematologic toxicity

Physicians predicted hematologic grade ≥ 3 toxicity better than overall toxicity. The toxicity risk increased with a higher physicians' judgment of risk (p = 0.068); see Figure 4-28. In univariate logistic regression, physicians estimated the differences in toxicity risk significantly, but only between the categories low and mid (low vs mid: OR 2.547, CI 1.056-6.144, p = 0.037; low vs high: OR 3.765, CI 0.894-15.851, p = 0.071). The ROC curve yielded an AUC of 0.621 (CI, 0.508-0.734; p = 0.039); see Figure 4-28.

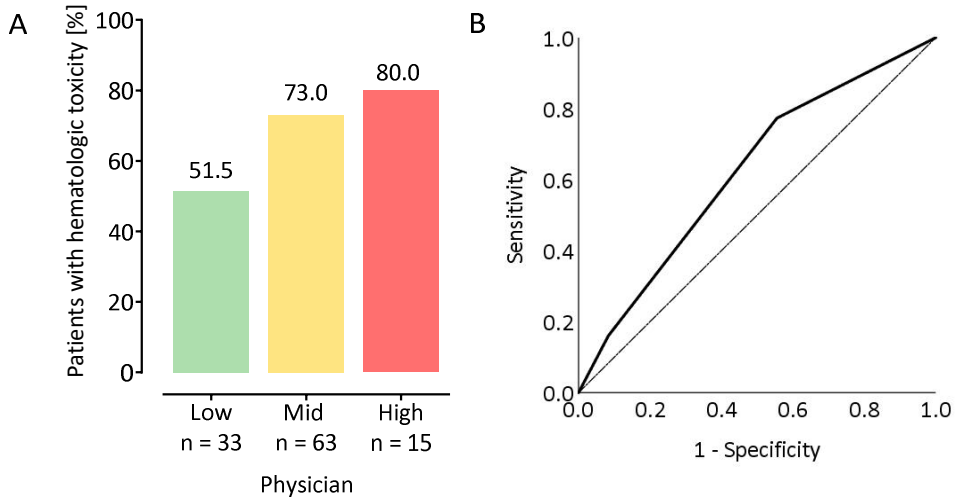


Figure 4-28 *(A) Proportion of patients with hematologic grade ≥ 3 toxicity per physicians' judgment of hematologic toxicity during therapy course; (B) ROC curve of the physicians' judgment for hematologic grade ≥ 3 toxicity during therapy course; solid line: physicians' judgment; thin line: line of identity; n = 111*

Nonhematologic toxicity

Nonhematologic grade ≥ 3 toxicity was not adequately predicted by physicians. The toxicity risk increased slightly with a higher physicians' judgment of risk (p = 0.562); see Figure 4-29. Physicians did not significantly predict nonhematologic toxicity in logistic regression (low vs mid: OR 1.257, CI 0.545-2.897, p = 0.592; low vs high: OR 2.040, CI 0.598-6.961, p = 0.255). The ROC-AUC demonstrated only low discrimination (ROC-AUC: 0.555, CI 0.447-0.664, p = 0.325), see Figure 4-29.

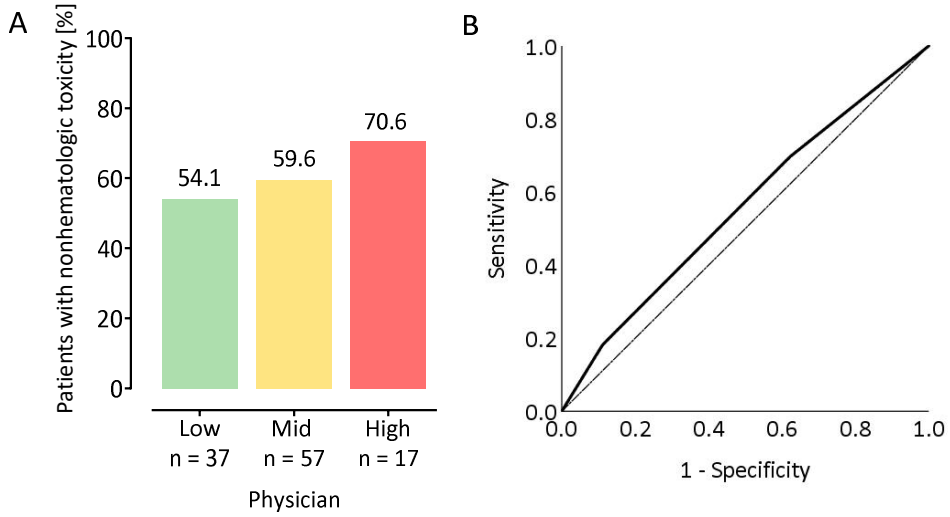


Figure 4-29 Proportion of patients with nonhematologic grade ≥ 3 toxicity per physicians' judgment of nonhematologic toxicity during therapy course (A); ROC curve of the physicians' judgment for nonhematologic grade ≥ 3 toxicity during therapy course (B); solid line: physicians' judgment; thin line: line of identity; n = 111

4.1.7.2 ECOG and age

Other commonly used predictors for the risk of severe toxicity are the ECOG performance status and chronological age. In our analysis for overall grade ≥ 3 toxicity, neither the ECOG performance status nor the age exhibited a sufficient predictive performance for justifying their use as predictor. The proportion of patients with overall grade ≥ 3 toxicity increased with higher ECOG category (p = 0.413, Fisher's exact test). For age, no consistent trend could be observed (p = 0.178, Fisher's exact test); see Figure 4-30. Neither the ECOG performance status nor the age predicted overall severe toxicity significantly in univariate logistic regression (ECOG: OR 1.712, CI 0.946-3.097, p = 0.075; age: OR 1.001, CI 0.900-1.112, p = 0.989). The discrimination was rather low for the ECOG performance status (ROC-AUC, 0.620; CI, 0.492-0.747; p = 0.088) and worthless for age (ROC-AUC, 0.460; CI, 0.339-0.580; p = 0.567), see Figure 4-31.

For hematologic toxicity, similar results were found; for nonhematologic toxicity, the ECOG performance status predicted toxicity adequately. Respective details are illustrated in Appendix D, Table D-13.

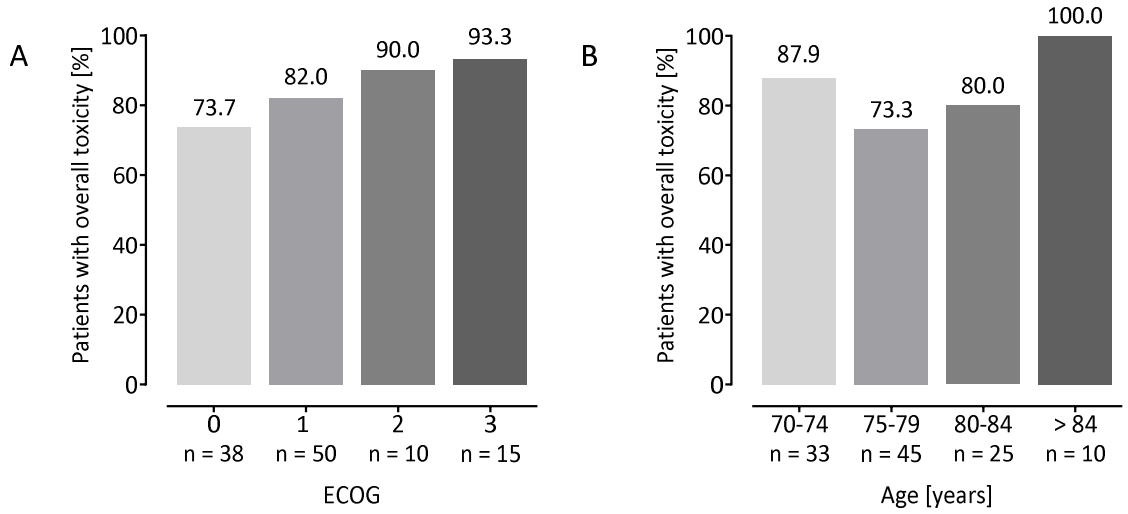


Figure 4-30 Proportion of patients with overall grade ≥ 3 toxicity during therapy course per ECOG (Eastern Cooperative Oncology Group) performance status (A) and per age (B); n = 113

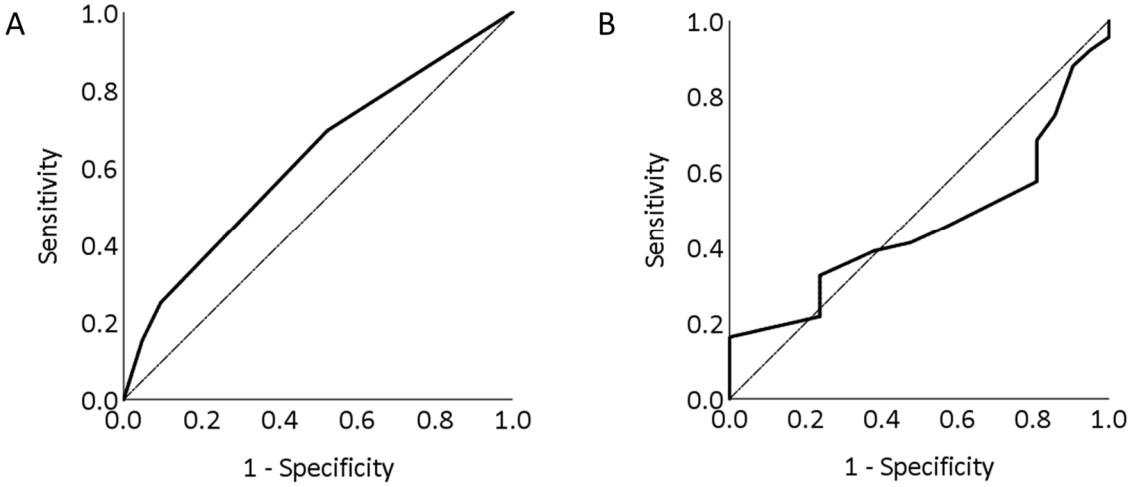


Figure 4-31 ROC curve of ECOG (Eastern Cooperative Oncology Group) performance status (A) and age (B) as continuous predictors for overall grade ≥ 3 toxicity during therapy course (n = 113); solid line: (A) ECOG, (B) age; thin line: line of identity

4.1.7.3 Individual CARG score items

From all items of the CARG score, only hemoglobin significantly predicted overall grade ≥ 3 toxicity during therapy course. Patients with a hemoglobin value below 10 g/dL (female) or below 11 g/dL (male) exhibited a four-fold higher risk of toxicity compared to patients with higher hemoglobin values (OR, 4.036; CI, 1.110-14.683; p = 0.034). All other items of the CARG score did not significantly predict overall toxicity when used separately from the CARG score, indicating that especially hemoglobin plays an important role for the predictive value of the CARG score in this patient cohort. Univariate logistic regression results of the CARG score items are displayed in Table 4-10. The distribution of grade ≥ 3 toxicity per score item is presented in Appendix D, Table D-1.

4.1.7.4 Individual CRASH score items

As only item from all CRASH score items, lactate dehydrogenase (LDH) significantly predicted overall severe toxicity during therapy course when analyzed separately from the score. Patients with higher LDH values featured an almost seven-fold risk compared to patients with lower LDH values (OR, 6.980; CI, 1.432-34.035; p = 0.016). The univariate logistic regression results of the CRASH score items are presented in Table 4-11.

4.1.7.5 Other patient- or cancer-related characteristics

Patients with targeted or immunotherapies instead of chemotherapy demonstrated a significantly lower risk of overall grade ≥ 3 toxicity (OR, 0.031; CI, 0.003-0.297; p = 0.003). All other patient or cancer-related characteristics did not show significant predictions. The results of the univariate logistic regression are illustrated in Table 4-12.

Table 4-10 Univariate logistic regression for the CARG score items regarding overall grade ≥ 3 toxicity during therapy course (n = 113); item categories are based on the original cut-offs in the CARG score; reference: in italic; nd, not determinable; CI, confidence interval; GI, gastrointestinal; GU, genitourinary

	Odds ratio (CI)	**P value**
Age *<72* vs ≥ 72	0.607 (0.071-5.218)	0.649
Cancer type *Other* vs GI/GU tumor	2.000 (0.423-9.457)	0.382
Dose *Standard* vs Reduced	1.705 (0.357-8.148)	0.504
Number of treatment agents *Monotherapy* vs Polytherapy	1.111 (0.284-4.349)	0.880
Hemoglobin [g/dL] *≥ 10* vs <10 (female), *≥ 11* vs <11 (male)	4.036 (1.110-14.683)	0.034
Creatinine clearance Jeliffe [mL/min] *≥ 34 mL/min* vs <34mL/min	nd	
Hearing abilities *Good/excellent* vs Fair/worse	2.991 (0.932-9.594)	0.065
Falls in past six months *0* vs ≥ 1	1.562 (0.416-5.860)	0.509
Medication intake *No assistance* vs Requires assistance	nd	
Limited in walking one block *Not limited at all* vs Limited	0.965 (0.374-2.494)	0.942
Decreased social activity because of health/emotional problems *A little or none of the time* vs Some, most, all of the time	3.200 (0.876-11.687)	0.078

Table 4-11 Univariate logistic regression for the CRASH score items regarding overall grade ≥ 3 toxicity during therapy course (n = 113); item categories are based on the original cut-offs in the CRASH score; reference: in italic; CI, confidence interval; ULN, upper limit of normal; IADL, instrumental activities of daily living; LDH, lactate dehydrogenase; ECOG, Eastern Cooperative Oncology Group; MMSE, Mini-Mental State Examination; MNA, Mini Nutritional Assessment; MAX2, chemotherapy toxicity index

	Odds ratio (CI)	**P value**
Diastolic blood pressure [mmHg] *≤ 72* vs > 72	1.292 (0.499-3.346)	0.598
IADL *26-29* vs 10-25	1.562 (0.416-5.860)	0.509
LDH [U/L] *≤ 0.74 x ULN* vs > 0.74 x ULN	6.980 (1.432-34.035)	0.016
ECOG performance status		
0	-	
1-2	1.786 (0.663-4.811)	0.252
3-4	5.000 (0.580-43.071)	0.143
MMSE *30* vs <30	2.229 (0.738-6.725)	0.155
MNA *28-30* vs <28	0.640 (0.133-3.077)	0.577
Therapy toxicity (MAX2)		
0	-	
1	1.233 (0.387-3.928)	0.723
2	2.467 (0.663-9.176)	0.178

Table 4-12 Univariate logistic regression for other potential predictors of overall grade ≥ 3 toxicity during therapy course (n = 113); categorial variables: reference in italic; if no categories are stated, the variable is treated as continuous; CI, confidence interval; ECOG, Eastern Cooperative Oncology Group

	Odds ratio (CI)	**P value**
Age	1.001 (0.900-1.112)	0.989
Sex *male* vs female	1.100 (0.426-2.841)	0.844
ECOG performance status	1.712 (0.946-3.097)	0.075
Charlson Comorbidity Index	0.979 (0.657-1.457)	0.916
Creatinine Clearance (Cockcroft-Gault) [mL/min]	0.999 (0.977-1.021)	0.921
Number of drugs before start of cancer treatment	1.145 (0.979-1.340)	0.090
Tumor type *Solid tumors* vs Hematological tumors	0.884 (0.342-2.286)	0.800
Treatment type		
Chemotherapy	-	
Targeted or immunotherapy	0.031 (0.003-0.297)	0.003
Combined chemotherapy and targeted or immunotherapy	0.732 (0.249-2.146)	0.569

Details regarding the univariate logistic regression for hematologic and nonhematologic grade ≥ 3 toxicity for different items are displayed in Appendix D, Table D-14 and Table D-15. For hematologic toxicity, hemoglobin (OR, 2.636; CI, 1.066-6.520; p = 0.036) and treatment type (targeted or immunotherapy vs chemotherapy: OR, 0.091; CI, 0.010-0.831; p = 0.034) were both significant predictors. Moreover, the MAX2 index predicted hematologic toxicity significantly. For nonhematologic toxicity, multiple variables showed significant predictions, also including hemoglobin and treatment type.

4.1.8 Prediction of severe symptom burden by onco-geriatric scores

The patients' total severe symptom burden (PRO-CTCAE ≥ 75%) was better predicted by the combined CRASH score than by the CARG score. In both scores, the proportion of patients with severe symptom burden increased with higher risk category (CARG: p = 0.459; combined CRASH: p = 0.356; Fisher's exact test); see Figure 4-32. In univariate logistic regression, only the combined CRASH score significantly predicted severe symptom burden. With each CRASH score point, the chance of developing severe symptom burden increased by 1.3-fold (OR, 1.344; CI, 1.028-1.757; p = 0.031). The CARG score denoted an OR of 1.119 (CI, 0.948-1.322; p = 0.184). The discriminative abilities of the combined CRASH score were superior to those of the CARG score (CRASH: ROC-AUC 0.66, CI 0.536-0.784, p = 0.024; CARG: ROC-AUC 0.581, CI 0.443-0.719, p = 0.253); see Figure 4-33.

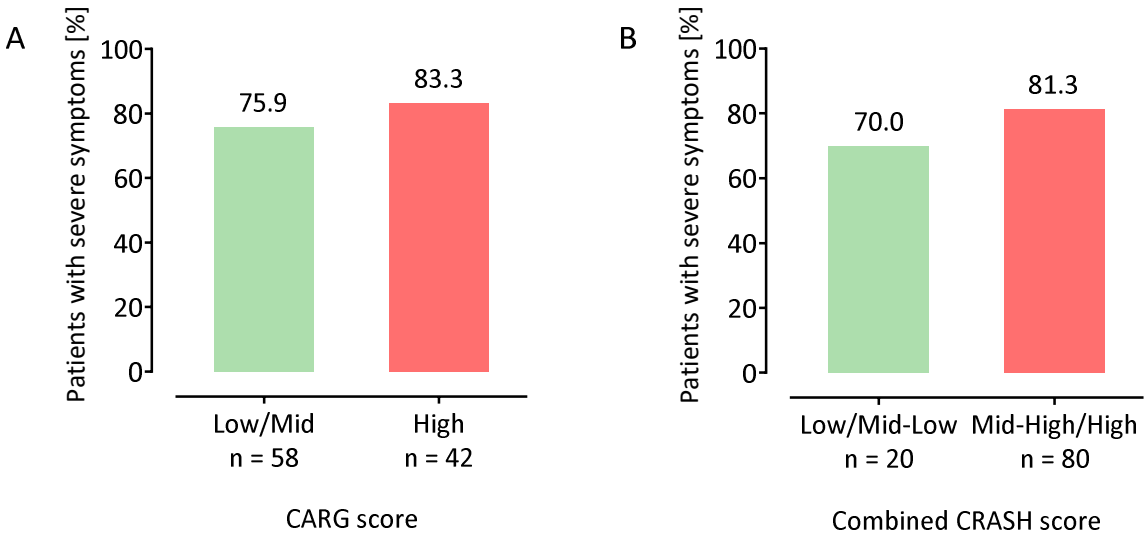


Figure 4-32 Proportion of patients with severe symptom burden (PRO-CTCAE ≥ 75) during therapy course per CARG score categories (A) and combined CRASH score categories (B); n = 100

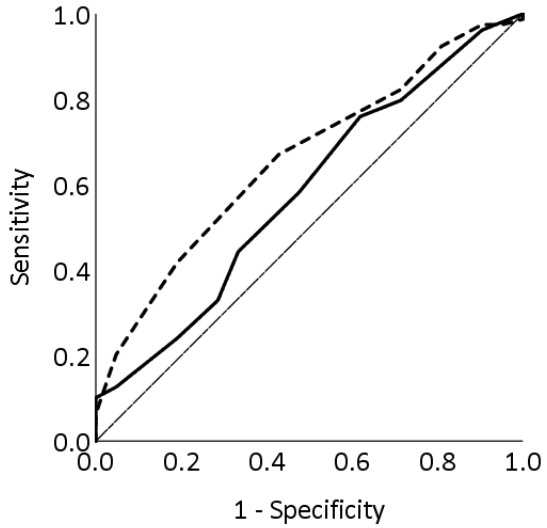


Figure 4-33 ROC curves of the CARG score and the combined CRASH score for severe symptom burden (PRO-CTCAE ≥ 75) during therapy course (n = 100); solid line: CARG score; dashed line: combined CRASH score; thin line: line of identity

4.1.9 Prediction of alterations of the planned treatment by onco-geriatric scores

For major alterations of the planned treatment (therapy discontinuations or changes), the CARG score as well as the combined CRASH score indicated a sufficient predictive performance. However, none of the onco-geriatric scores exhibited an adequate predictive performance for total alterations of planned treatment or minor alterations (delays/dose reductions). Corresponding details are shown in Appendix D, Table D-16 and Table D-17. The proportion of patients with therapy discontinuations or changes increased significantly with higher category in both scores (CARG: p = 0.002; combined CRASH: p = 0.031; Fisher's exact test). Figure 4-34 displays the proportion of patients with major alterations per score category. In univariate regression, both scores significantly predicted major alterations. With each increase of a CARG score unit, the risk of experiencing major alterations increased by 1.3-fold (OR, 1.261; CI, 1.102-1.443; p = 0.001). For the combined CRASH score, the risk increased with each unit by 1.5-fold (OR, 1.499; CI, 1.160-1.939; p = 0.002). Both scores denoted a similar ROC-AUC: A ROC-AUC of 0.696 (CI, 0.599-0.793; p = 0.001) was calculated for the CARG score and 0.682 (CI, 0.583-0.781; p = 0.001) for the combined CRASH score; see Figure 4-35. Both scores exhibited ROC-AUC values for the prediction of major treatment alterations close to those reached for the prediction of severe toxicity.

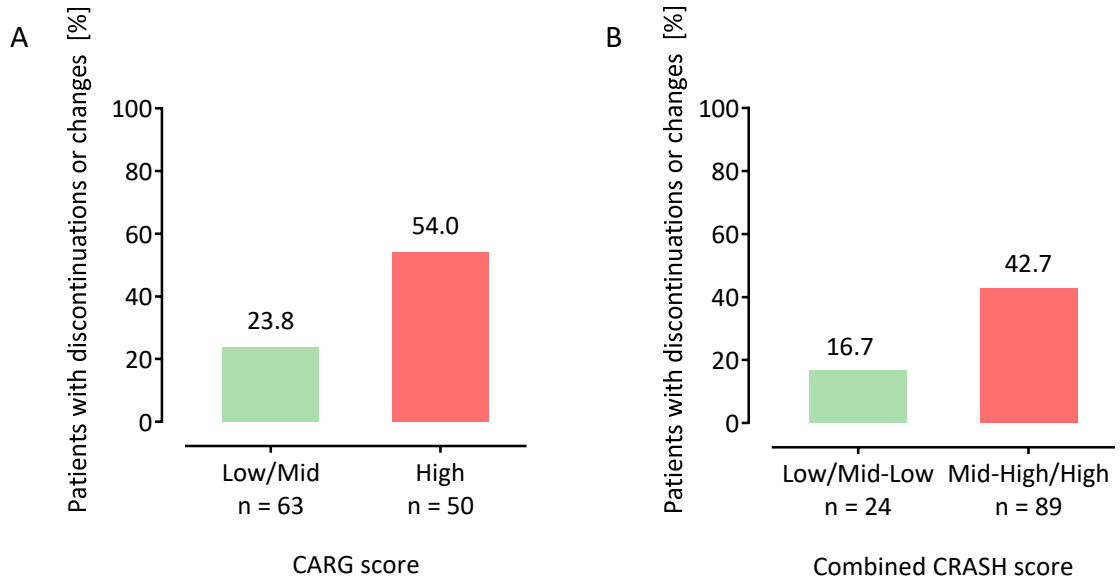


Figure 4-34 *Proportion of patients with major alterations of the planned treatment (discontinuations or changes) during therapy course per CARG score categories (A) and combined CRASH score categories (B); n = 113*

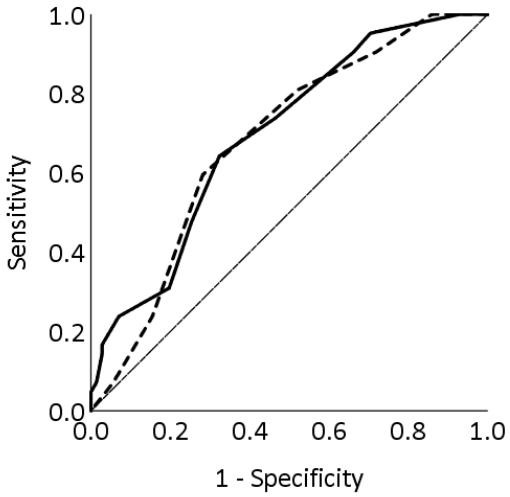
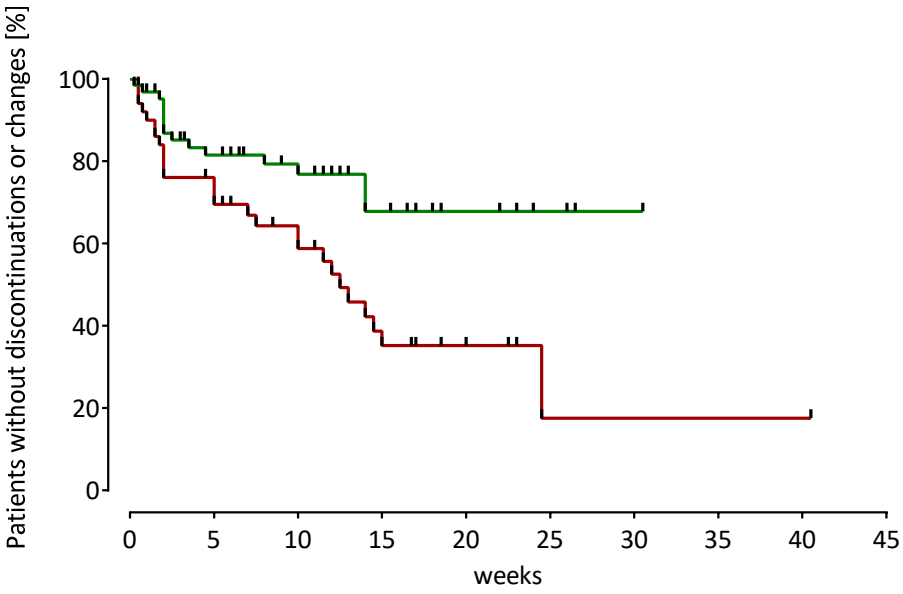


Figure 4-35 *ROC curves of the CARG and the combined CRASH score for major alterations of the planned treatment (discontinuations or changes); solid line: CARG score; dashed line: combined CRASH score; thin line: line of identity; n = 113*

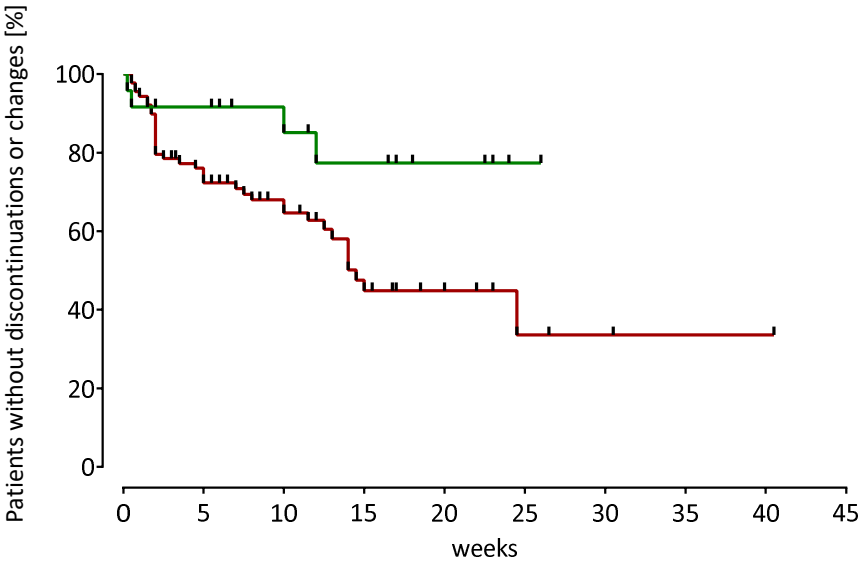
The onco-geriatric scores were not associated with the time until onset of total alterations or minor alterations (delays or dose reductions) of planned treatment; see Appendix D, Table D-18 and Table D-19. Those findings correspond to the results regarding the predictive performance for total and minor alterations.

However, both scores were associated with the time until onset of major alterations of the planned therapy (discontinuations or changes). A significant difference between low and high score categories was detected for the time to major alterations (CARG: p = 0.004; combined CRASH: p = 0.040; log-rank test). The Kaplan-Meier curves are illustrated in Figure 4-36 and Figure 4-37; for Kaplan-Meier estimates see Appendix D, Table D-20 to Table D-23. The Kaplan-Meier curves for the CRASH score crossed after about 2 weeks. However, this was not further taken into account since it was assumed to be caused by the documentation method: Since exact dates of the therapy alterations were not available, the middle of the cycle was considered as time point of alteration and thus time points in this short 1-2 week period might be imprecise. The median time to major alterations was not calculable for the low category in both scores because the event did not occur in at least half of patients. Regarding the high category, a median time to major alterations of 12.5 (CI, 9.573-15.427) weeks was estimated for the CARG score and a median time of 14.5 (CI, 12.456-16.544) weeks for the combined CRASH score. For both scores, the Cox regression demonstrated a higher risk for the high category (CARG: Hazard ratio (HR) 2.402, CI 1.277-4.516, p = 0.007; combined CRASH: HR 2.773, CI 0.989-7.781, p = 0.053). An exploratory analysis with cycles instead of weeks as time units yielded similar results.



Low	n	63	45	32	15	9	3	1	0	0
High	n	50	35	23	11	6	1	1	1	1

Figure 4-36 Kaplan-Meier analysis of the time to major alterations of the planned treatment (discontinuations or changes) dependent on the category of the CARG score (low-mid versus high); censored: low 76.2%; high 46.0%; green line = low CARG score; red line: high CARG score; vertical line: censored data; n = 113



Low	n	24	19	14	8	5	1	0	0	0
High	n	89	61	41	18	10	3	2	1	1

Figure 4-37 Kaplan-Meier analysis of the time to major alterations of the planned treatment (discontinuations or changes) dependent on the category of the combined CRASH score (low/mid-low versus mid-high/high); censored: low 83.3%; high 57.3%; green line: low CRASH score; red line: high CRASH score; vertical line: censored data

4.2 Medication risk analysis

4.2.1 Patient characteristics

For the analysis of medication risks, the data from the pilot and evaluation study were pooled, resulting in 136 included patients. The majority of patients originated from the evaluation study. An overview of the patients included in this analysis is given in Figure 4-38.

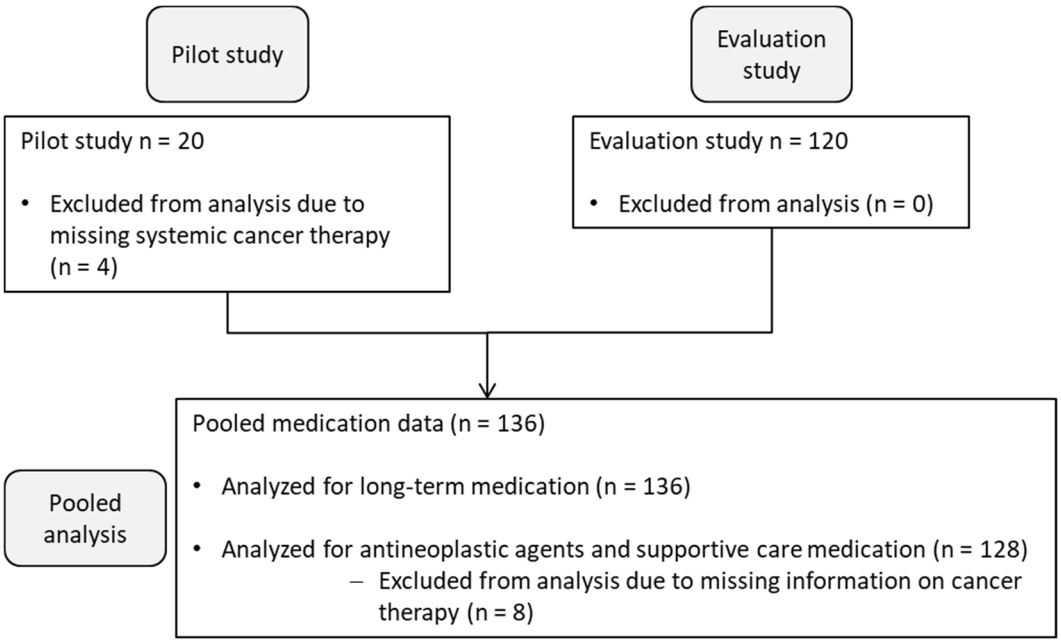


Figure 4-38 Flow chart of patients for medication risk analysis

The patient characteristics correspond mainly to those of the evaluation study (section 4.1.2); for details see Table 4-13.

*Table 4-13 Characteristics of the patients included into the medication risk analysis (n = 136); ECOG, Eastern Cooperative Oncology Group; * by body location according to the National Cancer Institute (NCI)*

Age [years]		
Mean (SD)	76.9 (4.53)	
Min-max	70-88	
Charlson Comorbidity Index		
Mean (SD)	1.05 (1.237)	
Min-max	0-7	
Creatinine Clearance (Cockcroft-Gault) [mL/min]		
Mean (SD)	67.2 (22.89)	
Min-max	10-131	
	n	**%**
Sex		
Female	68	50.0
Male	68	50.0
ECOG performance status		
Fully active (0)	45	33.1
Capable of all self-care (1-2)	74	54.4
Limited or no self-care (3-4)	17	12.5
Tumor entity *		
Respiratory	34	25.0
Hematological	57	41.9
Gynecological	5	3.7
Genitourinary	3	2.2
Unknown primary	4	2.9
Musculoskeletal	1	0.7
Digestive/gastrointestinal	16	11.8
Breast	13	9.6
Others	3	2.1
Relapse		
No	118	86.8
Yes	18	13.2

Table 4-13 (continued)

	n	%
Cancer stage		
I	7	5.1
II	11	8.1
III	31	22.8
IV	68	50.0
Missing	19	14.0
Treatment type		
Chemotherapy	81	59.6
Targeted or immunotherapy	9	6.6
Combined chemotherapy and targeted or immunotherapy	46	33.8

4.2.2 Long-term medication

Almost all patients were on long-term medication before the start of cancer therapy, only 8/136 (5.9%) patients did not take any regular long-term medication when being admitted to hospital. On average, patients took 5 drugs (SD, 3.5). The maximum number of drugs per patient was 18. Most drugs were only available on prescription (587/683); solely 96/683 drugs comprised over-the-counter (OTC) drugs. The most frequently used drug classes were antithrombotic agents (ATC Code B01; mostly acetylsalicylic acid, ASS), agents acting on the renin-angiotensin system (ATC Code C09; mostly ramipril), and diuretics (ATC Code C03; mostly hydrochlorothiazide). On the drug level, pantoprazole, L-thyroxine, and ASS were the drugs which patients received most often. Drug classes and individual drugs of patients' long-term medication are given in Table 4-14 and Table 4-15.

Table 4-14 Drug classes (ATC code level 2) of patients' long-term medication before start of cancer therapy (n = 136)

Drug class (ATC code level 2)	Number of drug prescriptions
Antithrombotic agents (B01)	70
Agents acting on the renin-angiotensin system (C09)	65
Diuretics (C03)	59
Beta blocking agents (C07)	50
Lipid modifying agents (C10)	48
Drugs for acid-related disorders (A02)	47
Thyroid therapy (H03)	46
Analgesics (N02)	32
Calcium channel blockers (C08)	30
Drugs used in diabetes (A10)	29
Others	207

Table 4-15 Individual drugs of patients' long-term medication before start of cancer therapy (n = 136); ASS, acetylsalicylic acid; HCT, hydrochlorothiazide

Drug	Number of patients	Proportion of patients with respective drug [%]
Pantoprazole	42	30.9
L-thyroxine	38	27.9
ASS	35	25.7
Simvastatin	31	22.8
HCT	29	21.3
Bisoprolol	22	16.2
Ramipril	22	16.2
Amlodipine	21	15.4
Metoprolol	20	14.7
Candesartan	17	12.5
Metamizole	17	12.5

4.2.2.1 Polymedication

More than half of patients exhibited polymedication (≥ 5 drugs) and approximately every 10th patient experienced hyperpolymedication (≥ 10 drugs). The prevalence of poly- and hyperpolymedication is illustrated in Figure 4-39.

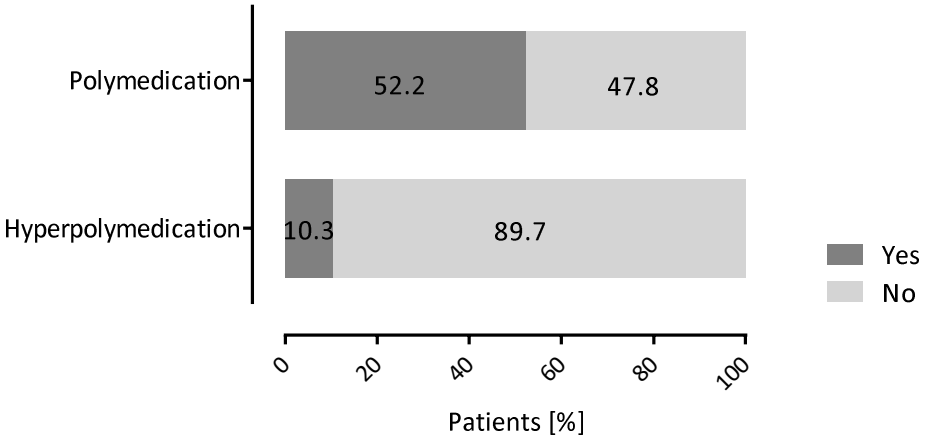


Figure 4-39 Patients with polymedication (≥ 5 drugs) and hyperpolymedication (≥ 10 drugs) before start of cancer therapy (n = 136)

4.2.2.2 Potentially inadequate medication

Patients took in median one (IQR, 1; range, 0-5) PIM drug. More than half of patients (52.9%) used at least one PIM drug before start of cancer therapy. By far the most frequent PIM drugs were drugs for acid-related disorders (ATC A02). Consistent with this finding, pantoprazole was the most frequently taken PIM drug (42/136 patients). Other commonly used PIM drug classes comprised drugs used in diabetes (ATC code A10; mostly sitagliptin), drugs for cardiac therapy (ATC code C01; mostly amiodarone), and calcium channel blockers (ATC code C08; mostly verapamil). An overview of the PIM drug classes and individual drugs is presented in Table 4-16 and Table 4-17.

Table 4-16 Prevalence of PIM drug classes (ATC code level 2) in long-term medication before start of cancer therapy (n = 136)

PIM drug class (ATC code level 2)	Number of drug prescriptions
Drugs for acid-related disorders (A02)	47
Drugs used in diabetes (A10)	10
Cardiac therapy (C01)	8
Calcium channel blockers (C08)	7
Psycholeptics (N05)	6
Psychoanaleptics (N06)	5
Antithrombotic agents (B01)	4
Anti-inflammatory and antirheumatic products (M01)	3
Beta blocking agents (C07)	3
Urologicals (G04)	3
Others	11

Table 4-17 Prevalence of individual PIM drugs in long-term medication before start of cancer therapy (n = 136)

PIM drug (ATC code)	Number of patients	Proportion of patients with respective drug [%]
Pantoprazole (A02BC02)	42	30.9
Sitagliptin (A10BH01)	8	5.9
Amiodarone (C01BD01)	4	2.9
Verapamil (C08DA01)	4	2.9
Rivaroxaban (B01AF01)	3	2.2
Omeprazole (A02BC01)	3	2.2
Amitriptyline (N06AA09)	3	2.2
Sotalol (C07AA07)	3	2.2
Diclofenac (M01AB05)	2	1.5
Diltiazem (C08DB01)	2	1.5
Methocarbamol (M03BA03)	2	1.5
Metoclopramide (A03FA01)	2	1.5
Pramipexole (N04BC05)	2	1.5
Trospium (G04BD09)	2	1.5

4.2.2.3 Relevant potential drug-drug interactions

Approximately one third of patients (30.9%) exhibited relevant potential drug-drug interactions (rPDDI) in long-term medication before start of cancer therapy. The median of rPDDI per patient was 0 (IQR, 1; range, 0-9). The majority of rPDDI was classified as “monitoring/modification needed” (67/71) according to the ABDA database. Only 4/71 rPDDI were categorized as “simultaneous usage not recommended”. No contraindications were observed. Mostly, rPDDI consisted of pharmacodynamic interactions (40/71); 21/71 rPDDI occurred due to pharmacokinetic reasons. In general, a variety of interaction types was observed; the most frequent rPDDI comprised “anti-diabetic drugs – corticosteroids”, “agents acting on the renin-angiotensin system – heparinoids” and “simvastatin – amlodipine”. The detected rPDDI are presented in Table 4-18.

The most frequent drug classes being involved in rPDDI were agents acting on the renin-angiotensin system (ATC code C09), beta blocking agents (ATC code C07), and antithrombotic agents (ATC code B01); see Table 4-19. According to the interaction propensity, the drug classes with the highest potential of provoking interactions were cardiac therapy (ATC code C01) and corticosteroids for systemic use (ATC code H02). Surprisingly, although being the drug class most frequently involved in interactions, agents acting on the renin-angiotensin system did not exhibit such a high potential of provoking rPDDI when prevalence was considered. The interaction propensity of different drug classes is depicted in Table 4-19.

Table 4-18 Types of rPDDI, observed in patients before start of cancer therapy; NSAID, nonsteroidal anti-inflammatory drugs (n = 136)

Type of interaction	Number of detected interactions
Anti-diabetic drugs – corticosteroids	8
Agents acting on the renin-angiotensin system – heparinoids	8
Simvastatin – amlodipine	8
Beta agonists – beta blocker	6
ACE inhibitors – allopurinol	5
Amiodarone – beta blockers	4
Thyroid hormones – polyvalent cations	4
Insulins – cardio selective beta blockers	3
NSAID – corticosteroids	3
Thiazide-diuretics – vitamin D and derivatives	3
Others	19

Table 4-19 Frequency of drug classes being involved in rPDDI and the respective interaction propensity (ratio of frequency as interaction partner per total frequency of a drug class); n = 136

Drug class (ATC code level 2)	Number of detected interactions	Interaction propensity
Agents acting on the renin-angiotensin system (C09)	17	0.26
Beta blocking agents (C07)	13	0.26
Antithrombotic agents (B01)	13	0.19
Corticosteroids for systemic use (H02)	12	0.75
Drugs used in diabetes (A10)	12	0.41
Lipid modifying agents (C10)	11	0.23
Diuretics (C03)	9	0.15
Cardiac therapy (C01)	9	1.0
Calcium channel blockers (C08)	8	0.27
Drugs for obstructive airway diseases (R03)	6	0.25
Others	32	-

4.2.3 Antineoplastic agents and supportive care medication

In total, 128 patients could be assessed after initiation of cancer therapy. After start of cancer therapy, patients received in median 6 (IQR, 2.25; range, 1-12) additional drugs. This comprised in median 2 (IQR, 1; range, 1-5) additional drugs for antineoplastic therapy and in median 4 (IQR, 2.25; range, 0-7) additional drugs for supportive therapy. Regarding antineoplastic agents, most frequently the drug classes plant alkaloids and other natural products (ATC code L01C, e.g. paclitaxel) as well as platinum compounds (ATC code L01XA, e.g. carboplatin) were prescribed in our cohort. Concerning supportive care medication, by far the most frequently used drug class was antiemetics and antinauseants (ATC code A04, e.g. ondansetron). G-CSF (granulocyte colony stimulating factor) was administered in 21/128 patients. Details regarding prevalence of antineoplastic agents and supportive care medication are presented in Appendix E, Table E-1 and Table E-2.

4.2.3.1 Potentially inadequate medication

After start of cancer therapy, 36.7% of patients received further PIM drugs being used more than once per cycle. In total, the median of additional PIM drugs for all patients was 0 (IQR, 1; range, 0-3). The most commonly used additional PIM drug was ranitidine. The prevalence values of additional PIM drugs after start of cancer therapy are listed in Table 4-20.

Table 4-20 Prevalence of additional PIM drugs after start of cancer therapy which are used more than once per cycle (n = 128)

PIM (ATC code)	Number of patients
Ranitidine (A02BA02)	32
Clemastine (R06AA04)	17
Proton-pump inhibitors (A02BC)	8
Dimetindene (R06AB03)	5
Alizapride (A03FA05)	2

4.2.3.2 Relevant potential drug-drug interactions

After start of cancer therapy, 29.7% of patients demonstrated further rPDDI between the long-term medication and the antineoplastic agents/supportive care medication which was used more than once per cycle. The median of additional rPDDI in all patients was 0 (IQR, 1; range, 0-3). The types of interactions were rather diverse. Most frequently, the rPDDI "NSAID – corticosteroids" and "cytotoxic agent – thiazide diuretic" were observed; see Table 4-21. The rPDDI were usually categorized as "monitoring/modification needed" by the ABDA database classification (Table 4-21). Three out of hundred twenty-eight (2.3%) patients exhibited contraindications; no patient experienced more than one contraindication. The most severe interaction types involved both QT prolonging agents. Most rPDDI consisted of pharmacodynamic interactions (30/50); changes in pharmacokinetics only rarely caused rPDDI (5/50).

*Table 4-21 Types of additional rPDDI between antineoplastic agents/supportive care medication (being used more than once per cycle) and the long-term medication after start of cancer therapy (n = 128), with prevalence and ABDA database classifications; NSAID, nonsteroidal anti-inflammatory drugs; * interaction was unintended in this case*

Type of interaction	Number of detected interactions	ABDA database classification
NSAID – corticosteroids	8	Monitoring/modification needed
Cytotoxic agent – thiazide diuretic	7	Monitoring/modification needed
Anti-diabetic drugs – corticosteroids	5	Monitoring/modification needed
ACE inhibitors – allopurinol	4	Monitoring/modification needed
Hyperkalemic drugs – trimethoprim	4	Monitoring/modification needed
QT prolonging drug – antidepressant	3	Simultaneous usage not recommended
QT prolonging drug – antiarrhythmic agent	3	Serious consequences possible – as precaution contraindicated
Loop diuretic – platinum compound	3	Monitoring/modification needed
Nitrogen mustard derivatives – allopurinol	3	Monitoring/modification needed
Fluoropyrimidine – folate *	2	Monitoring/modification needed
Others	8	-

Corticosteroids for systemic use (ATC code H02) was the drug class most frequently causing rPDDI. However, antibiotics (ATC code J01) were found to be the drug class with the highest interaction propensity (0.86). This was triggered by the numerous interactions between trimethoprim and ACE inhibitors. Respective details are illustrated in Table 4-22.

Table 4-22 Frequency of drug classes being involved in rPDDI between antineoplastic agents/supportive care medication (being used more than once per cycle) and the long-term medication after start of cancer therapy, with the interaction propensity (ratio of frequency as interaction partner per total frequency of a drug); n = 128

Drug class (ATC code level 2)	Number of detected interactions	Interaction propensity
Corticosteroids for systemic use (H02)	13	0.12
Diuretics (C03)	10	0.16
Antimetabolites (L01B)	9	0.30
Antithrombotic agents (B01)	8	0.11
Agents acting on the renin-angiotensin system (C09)	8	0.12
Antiemetics and antinauseants (A04)	8	0.07
Antigout preparations (M04)	8	0.13
Alkylating agents (L01A)	6	0.13
Drugs used in diabetes (A10)	6	0.21
Antibiotics (J01)	6	0.86
Others	18	-

4.2.4 Association of long-term medication with severe toxicity

For all 113 patients with a follow-up in the evaluation study, the associations of overall, hematologic, and nonhematologic grade ≥ 3 toxicity with medication risks were investigated by univariate logistic regression. For overall and hematologic toxicity, the occurrence of rPDDI was significantly associated with grade ≥ 3 toxicity: Patients with rPDDI exhibited an approximately 5-fold risk of developing overall toxicity (OR, 5.067; p = 0.036) and an approximately 4-fold risk of experiencing hematologic toxicity (OR, 3.949; p = 0.010).

However, the occurrence of rPDDI was not associated with nonhematologic toxicity. Instead, nonhematologic toxicity was significantly associated with the number of drugs per patient and the number of PIM drugs per patient. Corresponding details are displayed in Table 4-23. The distribution of toxicity regarding categorial variables is presented in Appendix E, Table E-3.

Table 4-23 Univariate logistic regression of overall, nonhematologic, and hematologic grade ≥ 3 toxicity during therapy course with risks in long-term medication (n = 113); reference: in italic; if no reference is given the variable was treated as continuous; Polymedication: ≥ 5 long-term drugs per patient

	Odds ratio (95% CI)	**P value**
Overall toxicity		
Number of drugs per patient	1.145 (0.979-1.340)	0.090
Patients *without* vs with polymedication	1.519 (0.584-3.954)	0.391
Number of PIM per patient	1.230 (0.710-2.131)	0.460
Patients *without* vs with PIM	1.310 (0.507-3.385)	0.578
Number of rPDDI per patient	3.843 (0.965-15.312)	0.056
Patients *without* vs with rPDDI	5.067 (1.109-23.140)	0.036
Hematologic toxicity		
Number of drugs per patient	1.037 (0.930-1.157)	0.511
Patients *without* vs with polymedication	1.173 (0.534-2.575)	0.691
Number of PIM per patient	0.908 (0.604-1.364)	0.642
Patients *without* vs with PIM	0.944 (0.430-2.077)	0.887
Number of rPDDI per patient	1.587 (0.899-2.803)	0.111
Patients *without* vs with rPDDI	3.949 (1.382-11.285)	0.010
Nonhematologic toxicity		
Number of drugs per patient	1.138 (1.014-1.277)	0.029
Patients *without* vs with polymedication	1.468 (0.691-3.121)	0.318
Number of PIM per patient	1.675 (1.051-2.669)	0.030
Patients *without* vs with PIM	1.926 (0.900-4.120)	0.091
Number of rPDDI per patient	1.591 (0.952-2.658)	0.076
Patients *without* vs with rPDDI	1.663 (0.715-3.870)	0.238

5 Discussion

5.1 Evaluation of onco-geriatric scores

For predicting therapy-associated toxicity in older cancer patients, the ASCO guideline recommends the CARG and the CRASH score. Nevertheless, it remains unclear which score to prefer. This study observed that the CARG and the CRASH score exhibited a similar predictive performance for severe overall toxicity. However, for predicting severe hematologic toxicity, the hematologic CRASH score should be preferred to the CARG score. Both scores performed better for overall toxicity than physicians' judgment alone, emphasizing the importance of onco-geriatric scores in complementing the clinical prediction of therapy-associated toxicity.

5.1.1 Study set-up

Study design

This is the first study directly comparing the CARG and the CRASH score in a clinical routine setting. Moreover, this study fills a gap in knowledge, being the first CARG score study which includes patients with hematologic malignancies. For the CRASH score, this is the first study investigating its predictive performance in a patient cohort different from the development study, hence evaluating its external validity. Furthermore, no previous study has investigated the scores in a cohort of patients with targeted or immunotherapies. Thus, this research enhances the applicability of these onco-geriatric scores in the current clinical routine where therapies like targeted therapy or immunotherapy play an increasing role. Another strength of this study is that its feasibility was verified by conducting a pilot study with 20 patients (in parts presented in [78]).

A limitation of this study is the retrospective collection of follow-up data from medical records. Toxicity might not have been documented thoroughly. However, this study focuses on grade ≥ 3 toxicity which is likely to be adequately documented since mostly evoking clinical measures. Furthermore, if blood controls were performed at general practitioners and oncology practices, those were contacted for the collection of medical data. This study therefore considers the patients' entire therapy course by including all relevant laboratory

data, also the nadir of blood values, into the analysis. Moreover, the data collection was standardized via a toxicity documentation form (Appendix C) and was always carried out through the same researcher to ensure consistent data quality.

The relatively small sample size and single-center design might limit the generalizability of the results. Due to lacking data, no sample size calculation could be conducted. Hence, we aimed at recruiting as many patients as possible in the given time frame of the project. Furthermore, the inclusion of patient data from external oncology practices might bear potential for bias due to different documentation standards. However, most patients (77.0%) were treated in the Johanniter Hospital Bonn, assuring a consistent documentation. Moreover, the inclusion of patients partly continuing treatment in an outpatient setting has the benefit of avoiding a selection bias since allowing for a study cohort closer to clinical routine.

Patient population

Broad inclusion criteria and only limited exclusion criteria were defined for this evaluation study because the onco-geriatric scores were designed to comprise a broad prediction scope across different tumor entities and cancer therapies. Furthermore, a heterogeneous study population better represents the real population of a hospital. Thus, the results are more eligible to be transferred into clinical routine and comprise a higher external validity. For enhancing applicability in current daily routine, not only chemotherapy but also targeted or immunotherapies were included, as well as inpatient and outpatient patients.

The onco-geriatric scores are particularly important for frail patients. By setting the inclusion criterion to ≥ 70 years and conducting recruitment in an inpatient setting, we targeted at enrolling this frail patient population. However, eventually, our study cohort consisted of mainly fit patients (ECOG 0-2, 87.5%), thus not representing a typical cohort of geriatric patients. This was also the case in the development studies of the CARG and the CRASH score (CARG: Karnofsky performance status > 70, 80%; CRASH: ECOG 0-2, 97% [67, 68]) and might be caused by the common selection bias of non-frail patients in studies. This seems likely since in this study, a frequent reason for exclusion comprised physical constraints.

Different eligibility criteria were used in the original development studies of the CARG score and the CRASH score [67, 68]. Due to the comparative character of our study, we considered both eligibility criteria as much as possible when designing this evaluation study. However,

being based on both development studies, our study featured different eligibility criteria than previous development and validation studies which only investigated one score. A comparison of the eligibility criteria of the score development studies with those of this evaluation study is shown in Table 5-1.

Table 5-1 Comparison of eligibility criteria of the CARG and CRASH score development studies with this evaluation study; AML, acute myeloid leukemia

	CARG score [67]	**CRASH score** [68]	**Evaluation study**
Age	≥ 65 years	≥ 70 years	≥ 70 years
Tumor entity	Solid tumors	Solid tumors and hematological malignancies (except AML)	Solid tumors and hematological malignances
Additional therapy	Radiotherapy allowed	No radiotherapy allowed	Radiotherapy allowed
Systemic cancer therapy	Only chemotherapy	Chemotherapy and combinations with targeted therapies	Chemotherapy, combinations with targeted or immunotherapies, and targeted or immunotherapies only
Previous therapy	Starting new chemotherapy regimen	Starting new line of chemotherapy (first-line to fourth-line)	Only first-line therapy

In consequence, the study population of the score development studies and this evaluation study differed: The study cohort of the CARG score development was younger than our cohort (CARG: mean 73 years, this study: mean 77.2 years) since Hurria et al. enrolled patients ≥ 65 years; in the CRASH score development study, the mean age was 75.5 years, also being slightly lower than in our study cohort [67, 68]. Gender was equally distributed in our study as well as in both score development studies (CARG: female 56%, CRASH: female 50.4%, this study: 50.0%). The CARG score development study only considered patients with solid tumors, whereas the CRASH score and this evaluation study also included hematologic tumors. Lymphoma were the most frequent hematologic malignancies in our study (27.5%) as well as in the CRASH development study (non-Hodgkin lymphoma 15.1%). However, our study included more patients with hematologic malignancies than the CRASH development study.

In all studies, lung cancer indicated the highest prevalence among solid tumors (CARG: 29%, CRASH: 20.8%, this study: 24.2%). Furthermore, patients experienced in all studies later cancer stages (≥ stage III: CARG 83%, CRASH 78.9%, this study 72.5%). In the CARG score development, no concomitant radiotherapy was allowed; however, in the CRASH score development 18.5% of patients and in our study 30.8% of patients were additionally treated with radiation therapy. Whereas the development studies of the CARG and CRASH score only focused on chemotherapy regimens, our study also comprised modern therapies. However, patients were rarely exclusively treated with targeted or immunotherapy regimens in our study (5.8%). Instead, combinations of targeted or immunotherapy with chemotherapy were more common (34.2%). Thus, our results are probably of less validity for exclusively targeted or immunotherapy regimens but rather valid for combinations of modern therapies with chemotherapy.

5.1.2 Risk assessment

Onco-geriatric scores

The time between the performance of the scores and the start of cancer therapy was only short, being in median 1 day. Hence, it can be assumed that the score results remained valid at start of therapy. However, some geriatric assessment variables like the MMSE are generally dependent on the daily condition of the patients. Also, some patients might have under- or overestimated their own performance for e.g. IADL. In general, wrong self-evaluation of patients is an inherent limitation of geriatric assessment variables. In order to ensure consistent conditions for patients, the different geriatric assessment instruments were always conducted in the same order and with the same interviewer. The questions of the CARG score were translated into German by our research group and asked orally in a patient interview. For allowing the use of a self-administered CARG questionnaire in the future, it would be an interesting field of further research to linguistically validate those questions. Regarding the CRASH score, it might be interesting to linguistically validate the IADL, since no validated German version exists.

The CARG score results in this evaluation study indicated a mid risk as most frequent prediction, consistent with the results of the development study (CARG study: 48.9% vs this

study: 50.8%). However, compared to the development study, much less patients were categorized as low risk (CARG study: 27.6% vs this study: 5.8%). The high category was predicted more often in this evaluation study than in the original development study (CARG study: 23.5% vs this study: 43.3%). Those differences may originate from different eligibility criteria: For instance, due to the inclusion criterion ≥ 70 years in this evaluation study compared to ≥ 65 years in the CARG study, only 7.5% of our patient cohort comprised the low category for age, compared to 46% in the CARG development cohort (< 72 years: 0 score points assigned vs ≥ 72: 2 score points) [67].

Regarding the combined CRASH score in this evaluation study, most patients were categorized as mid to high risk for overall toxicity and only few patients as low risk (mid-high: 60.0%; low: 2.5%); similar results were obtained for the hematologic and nonhematologic CRASH score. The distribution of the CRASH score results was not described in the development cohort. In literature, only one study was found which applied the CRASH score [127, 128]. In this study with non-Hodgkin lymphoma patients, the combined CRASH score predicted a higher risk (high or mid-high) in about half of patients [127], being slightly less than in our cohort. The results for the CRASH score items did not differ from the ones observed during the development study (e.g. MMSE: this study median 28 vs CRASH study median 28; MNA: this study median 24 vs CRASH study median 25; IADL: this study median 28 vs CRASH study median 28) [68].

Physicians

In contrast to the onco-geriatric scores, mainly predicting mid-high risk, physicians mainly expected a mid or low risk for their patients. In general, physicians were thus more optimistic regarding the tolerability of treatment. These results are in line with the results of a study by Moth et al. where physicians primarily predicted medium and low toxicity risk as well (low: 24%, mid: 63%, high: 12%) [73]. However, instead of asking to estimate an exact percentage of toxicity as Moth et al. [73], this study asked physicians to estimate risks as low, mid, or high. Since physicians are not trained for such assessments, we did not expect them to be able to estimate risks in such a detailed manner as percentages. However, physicians could have had different perceptions about the meaning of “low, mid, or high” which might have led to a bias of judgments. In addition, the physicians’ lack of training may have caused a subjectivity in the risk prediction and in consequence an incoherence in the risk assessment. Also, we asked the treating physicians who were mostly assistant physicians with limited experience. However,

those physicians were the ones directly treating the patients and hence were deemed to know the patient's health condition best.

5.1.3 Outcome

Toxicity

In total, 113 patients could be considered for outcome analysis. Seven patients were not included because the follow-up was incomplete, bearing the risk of misclassification which could lead to either under- or overestimation of the predictive performance. The loss to follow-up occurred early in all cases (mostly during or after the first cycle). Although our study showed that severe toxicity occurred frequently at start of therapy, the toxicity outcome of those patients could not be determined with adequate accuracy. The amount of missing laboratory data was not substantial in the follow-up (median 2.8%) and thus did not constitute a limitation in this study.

This evaluation study ended the follow-up after six cycles instead of observing the therapy course until the end of therapy like in the CARG development study, or following until a maximum of six months like in the CRASH development study [67, 68]. This might have prevented capturing toxic symptoms occurring later during therapy course. However, for 58.4% of patients in this evaluation study, the follow-up ended before six cycles. Moreover, results of the CRASH development study as well as our results indicated that severe toxicity mostly occurs at start of therapy: In this evaluation study, in 69.0% of patients, toxicity already occurred during the first cycle of therapy course; the median time until the first occurrence of toxicity was 22 days in the CRASH study [68] and 2 weeks in this study. The duration of cycles differed for each therapy regimen, resulting in different total lengths of follow-up. However, cycles were chosen as time frame for follow-up since this generally better reflects the therapy course than a certain number of months. This evaluation study did not further consider follow-ups when patients completely changed therapy regimen or experienced a dose reduction of ≥ 50%. In the development studies, the follow-up was not discontinued in those cases. However, since the therapy regimen and dosage are included in the score calculation, and thus in the score predictions, the therapy course after change of regimen or dosage would not reflect the

previous predictions anymore. Nevertheless, the regimen or dosage was mostly changed due to severe toxicity, implying that severe toxicity, the primary endpoint, had already occurred.

Overall, severe toxicity incidence was by far higher in our study (81.4%) than in the CARG score development study (53%) as well as the CRASH development study (64%) [67, 68]. Also, severe hematologic toxicity (CARG study 26%, CRASH study 32% vs this study 67.3%) and nonhematologic toxicity (CARG study 43%, CRASH study 56% vs this study 59.3%) was more frequent than during the score development studies [67, 68]. Furthermore, in our study, hematologic toxicity occurred more frequently than nonhematologic toxicity which was the opposite in the development studies. These findings might be explained by the different eligibility criteria: Contrary to the CARG score development study, this study allowed concomitant radiation therapy and enrolled older patients, patients with hematologic malignances, and patients in an inpatient setting. In contrast to the CRASH score development study, this study included more patients with hematologic malignancies, being more susceptible to severe hematologic toxicity. Also, differences in the endpoint collection could have caused a higher toxicity incidence: For the CARG score development, toxicity was only considered if, after being reviewed by two physicians, the toxicity was regarded as therapy-associated [67]. In the CRASH score development, only grade 4 but not grade 3 hematologic toxicity was defined as severe toxicity. If our study only considers grade 4 hematologic and grade ≥ 3 nonhematologic toxicity (similar to the CRASH score development), a prevalence of 70% would be observed for severe toxicity, being closer to the 64% observed during the CRASH score development [68]. Furthermore, the CARG score development study only considered blood values if they were measured on the day of scheduled chemotherapy or at emergency visits; the CRASH score development study also accounted for weekly complete blood counts. To capture the whole picture, this evaluation study considered all blood values. This also might have led to a higher toxicity incidence than during the CARG development study [67].

The distribution of toxicity types was similar in this evaluation study and in the CARG development study [67]. Neutropenia and leukopenia were the most frequent hematologic toxicity types in this evaluation study as well as in the CARG development study; for nonhematologic toxicity, fatigue and infections were frequent toxicity types in both studies [67]. The time to occurrence of severe toxicity was comparable with the CRASH development study where the time to first severe toxicity was 22 days [68]; in this evaluation study, the first

severe toxicity occurred after 2 weeks. Our study indicated that the majority of overall severe toxicity occurred at start of treatment (69.0% of patients showed toxicity at start of therapy vs 81.4% of patients during complete therapy course). This result is in line with a study by Extermann et al., observing that 46% of patients experienced the first severe toxicity during the first cycle and less patients in the following cycles (second cycle: 24%, third cycle: 17%) [95].

Patient-reported symptom burden

Evaluating patient-reported symptoms in addition to health care providers‘ information is essential for obtaining the complete picture of therapy tolerability [97]. It is a strength of this study that patient-reported outcomes were collected as well.

The PRO-CTCAE questions were asked orally in person or via telephone. Orally, the PRO-CTCAE was validated for interactive voice response systems [102] but not for patient interviews. Results may have been influenced by the fact that patients talked to a real person, for instance some symptoms might have been embarrassing for them to admit. However, since a voice response system was not available due to financial and logistic constraints, this mode of application was the only one feasible in our study in order to include data from patients not continuing treatment in the Johanniter Hospital Bonn.

Since the recall time of the PRO-CTCAE is seven days [94], we set the time frame for contacting patients to one to two weeks after the start of each cycle. As the patients were not always reachable via telephone or the patients experienced unforeseen therapy delays or changes, this time frame was slightly exceeded in some cases. However, since the length of the cycles differed and the chemotherapy varied from one to several administration days per cycle, this might not have influenced results to a large extent in this heterogeneous setting. Ideally, patients should be observed each week. However, since patient burden should be kept low due to their high age of patients, that did not seem feasible in this study.

A high percentage of patients was missing for follow-up. This is largely caused by the fact that the older study cohort did not tolerate cancer therapy well – many patients refused to be interviewed due to bad health condition. This implies a bias since patients with bad tolerability could not be asked for reporting their symptoms. However, this bias is an inherent challenge of patient-reported symptoms.

Agreement of toxicity and patient-reported symptom burden

When comparing severe patient-reported symptom burden (≥ 75% PRO-CTCAE) with severe physician-reported toxicity (grade ≥ 3 CTCAE), the weighted Kappa results generally indicated a low agreement. This is consistent with literature: A systematic review by Atkinson et al. summarized that the agreement between CTCAE and PRO-CTCAE was found to be low to – at best – moderate [129]. Furthermore, this study found that severe symptoms were reported more frequently by patients than by physicians. In line with this result, several studies in literature indicate that subjective toxic symptoms are likely of being underreported by physicians [96, 130, 131]. Those results were also observed in older cancer patients: Moon et al. found that physicians underreported toxicity in older head and neck or lung cancer patients during curative radiotherapy [132].

In this analysis, a PRO-CTCAE of 75% was used as the cut-off for severe toxicity. However, another cut-off may better correspond to severe toxicity being defined as CTCAE grade ≥ 3. The choice of a different cut-off value might be an interesting field of further investigation. Moreover, physicians' reporting of toxicity might have been underestimated due to the retrospective collection of data in the medical records. However, our results were consistent with previous studies, supporting the robustness of results.

Nevertheless, it should be kept in mind that the CTCAE and PRO-CTCAE were not developed for replacing but, in contrast, for complementing each other. The physicians' perspective and patients' perspective comprise both essential, meaningful information [97]. Basch et al. showed that the CTCAE better predict unfavorable clinical events, whereas patient-reported outcomes better predict the daily health status [97].

Alterations of the planned treatment

A high proportion of patients experienced alterations of the planned treatment in this evaluation study. Twenty-seven percent of patients discontinued therapy due to toxicity in our study, which was comparable to the results of the CRASH development study, where 23.4% discontinued therapy due to toxicity [68]. The percentage of patients with dose delays corresponded to the results of the CARG score development study (this study: 33.6% vs CARG study: 31% [67]). Dose reductions were less prevalent in our study than in the CARG development study (this study: 17.7% vs CARG study: 31% [67]). In a study by Wildes et al.,

16.9% of patients did not complete the planned number of therapy cycles due to toxicity, which is less than in this study where 37.2% of patients discontinued or changed their regimen [133]. This might be explained by differences in the study cohort: Wildes et al. enrolled younger patients and overall toxicity incidence only reached 41%. A Korean study observed that 40% of older patients discontinued cancer therapy during or after first-line therapy because of death or deteriorated health condition [134].

5.1.4 Relationship between risk assessments

Only poor agreement between the CARG and the CRASH score prediction was found. This indicates that both scores predict different risks for patients and are not interchangeable. Despite different predictions, the predictive performance of the onco-geriatric scores was similar. That finding suggests that the scores predict well for different patients, predicting complementary risks.

Moreover, physicians' judgments were not consistent with the predictions of the scores. Similar results were found in a study by Moth et al. and Alibhai et al. where physicians' judgments were not correlated with the CARG score results (Moth et al.: $r = -0.03$; Alibhai et al.: $r < 0.3$) [72, 73]. Also, Nishijima et al. found little agreement between physicians' treatment decision "reduced vs standard therapy" and "low vs high CARG score category" (Kappa value 0.14) [74]. Clinical judgment might thus be complemented by the scores.

5.1.5 Predictive performance of the onco-geriatric scores

Predictive performance of the onco-geriatric scores for toxicity during therapy course

The CARG and the CRASH score exhibited a similar adequate predictive performance for severe overall and nonhematologic toxicity. For predicting severe hematologic toxicity, the hematologic CRASH score should be preferred to the CARG score.

In general, the more a ROC-AUC approaches 1, the better the discrimination [115]. For overall toxicity, the CARG score and the combined CRASH score were relatively far from 1, showing a ROC-AUC of 0.681 and 0.650, respectively. However, these ROC-AUC results were close to

those reached in the validation study of the CARG score (0.65) and the combined CRASH score (0.64) [68, 83]. Those, in general, rather low ROC-AUC values are consistent with the finding that a geriatric assessment predicts therapy-related toxicity rather moderately at the individual patient level [20]. This might be caused by the high number of factors influencing the individual toxicity risk [20]. The proportion of patients with overall toxicity increased with higher score category in our study. However, this did not reach statistical significance which might be explained by the moderate sample size. For hematologic toxicity, the hematologic CRASH score performed better than the CARG score (AUC-ROC 0.665 vs 0.564, respectively) which was consistent with our expectations since the hematologic CRASH score was developed for this type of toxicity [68].

The predictive value of the CARG score has been assessed in different studies. In a recent study by Moth et al., the CARG score did not show a predictive value for patients with solid tumors (AUC-ROC 0.52; OR 1.04; p = 0.54; no increase of toxicity incidence with CARG score risk category) [73]. Contrary to our study, Moth et al. did not include hematologic malignancies. For prostate cancer patients, Alibhai et al. could not demonstrate a predictive value of the CARG score either (AUC-ROC 0.54; OR 1.09; p = 0.58). Toxicity increased with the CARG score category but not significantly (p = 0.65) [72]. However, the study of Alibhai et al. was limited by a relatively small sample size of 46 patients. In lung cancer patients, Nie et al. observed that toxicity incidence increased significantly with higher CARG risk category [71]. Of note, this study amended the CARG score by deleting the item tumor type as all patients were lung cancer patients, and by applying new cut-off values for risk categories. Nishijima et al. found that patients with a CARG score of ≥ 10 experienced toxicity more often than patients with a CARG score of < 10, being similar to our results [74].

For adjusting the scores to our patient cohort, the Youden index suggested that the cut-off values of the scores might be slightly changed. Regarding overall toxicity, the CARG score would exhibit an ideal cut-off between the mid and high category at ≥ 9, thus one point lower than the original cut-off at ≥ 10 [67]. For the combined CRASH score, a cut-off at ≥ 8 instead of ≥ 10 [68], might be ideal. The cut-offs for differentiating between the categories low and mid were difficult to assess due to the small patient numbers in the low category. The sensitivity and specificity for the optimal cut-offs were moderate, ranging between 0.6 and 0.7. As discussed in section 5.1.3, more toxicity occurred in this evaluation study than during

the development studies of the onco-geriatric scores. Hence, only the categorization of the scores (low-mid-high) can be deemed appropriate, but not the exact percentages of toxicity risk per score value being derived from the development studies. Problems with calibration-in-the-large (systematic under- or overprediction in external cohorts) occur frequently in predictive models [135].

Since both scores indicated a similar predictive performance for overall and nonhematologic toxicity, other factors are decisive for determining which score to use. Since the CARG is quicker and easier to use compared to the CRASH score [62], the CARG score may be preferred in busy daily routine. However, if estimates for hematologic toxicity are needed, the CRASH score is preferable. The CRASH score could also be incorporated in a full CGA as it already comprises various detailed geriatric assessment tools [70].

The impact and implementation of a risk prediction model are two essential issues during the evaluation process [136, 137]. Models which do not change behavior are not useful [138]. However, those steps are not yet fully assessed for the CARG score and the CRASH score. Thus, further studies should be conducted for evaluating their use and value in clinical routine.

The applicability of the onco-geriatric scores is broad: Clinical trials may use these scores as tools for risk stratification. In clinical routine, the onco-geriatric scores could be applied for a shared decision-making since toxicity is an important aspect for the therapy decision of patients [66]. The prediction of toxicity might also allow for better therapy individualization, possibly reducing under- and overtreatment of older cancer patients. However, the scores may only support and should not fully substitute the clinical decision-making process. Clinical decisions should consider more than only toxicity, for instance life expectancy or patient preferences. For judging those areas, other instruments exist: For example, life expectancy can be evaluated by the Onco-MPI (Oncological-Multidimensional Prognostic Index) [139]. A study by Moth et al. found that 83.3% (25/30) of physicians judged the CARG score as useful – but for most physicians, the CARG score did not influence treatment decisions. Reasons for that were for example missing familiarity with the score or problems translating the score results into a modification of therapy [140]. Furthermore, the scores might be used in other fields. The CARG score has also shown to be predictive for frailty [141] and hospitalization [74]. Also, the CARG score and the CRASH score were deemed as useful screening tools for detecting geriatric patients in need for a CGA [142].

Predictive performance of the onco-geriatric scores for toxicity at start of therapy

Since most of the toxicity occurred when therapy started, it was assumed that the onco-geriatric scores may also predict toxicity at start of therapy. For overall and nonhematologic toxicity, this was also the case in our study – but not for hematologic toxicity. The hematologic CRASH did not demonstrate any superiority compared to the CARG score for predictions at start of therapy. This might be due to the fact that a higher proportion of hematologic toxicity occurred later during therapy course than nonhematologic toxicity.

Time-related predictions

Both scores generally indicated a faster onset of toxicity in the high category compared to the low category. Thus, the scores may not only give information about *whether* severe toxicity occurs but also about *when*. This might be critical for patients because longer times until onset of toxicity imply a longer preservation of an adequate quality of life and a prevention of early toxicity-related therapy modifications.

The CARG score categories low vs high were significantly associated with the time to severe overall, hematologic, and nonhematologic toxicity. The hematologic CRASH score indicated a significantly faster onset of hematologic toxicity in the high category as well. In line with the results regarding the predictive performance for hematologic toxicity at start of therapy, the hematologic CRASH score was not superior to the CARG score. For overall and nonhematologic toxicity, however, the hazard ratio between the CRASH score categories difference in time until onset of severe toxicity was not significant. The lack of significance might be due to the uneven distribution of patients in the categories: The low CRASH category consisted of only 24 patients, compared to 89 patients in the high category. In contrast to this, the CARG score split the patient cohort more evenly (low: 63 patients, high: 50 patients). In general, only the first part of the Kaplan-Meier plot could be reasonably assessed due to the low patient numbers in time points after approximately 10 weeks. Therefore, the crossing of the CRASH score curves towards the end of the follow-up was not considered as violation of the Cox regression assumption. Further research should be conducted to verify this exploratory Kaplan-Meier analysis in a larger patient cohort. Regarding data collection, it must be considered that the middle of the cycle was assumed as the time point of onset of toxicity since the actual day of onset was mostly not available. Thus, the results could be diluted to a

certain extent. However, since the scores exhibited such a clear difference in the time until onset of toxicity, this is unlikely to have influenced the overall conclusion.

Predictive performance of the onco-geriatric scores for symptom burden

Patients' total severe symptom burden (PRO-CTCAE $\geq$ 75%) was better predicted by the combined CRASH score than by the CARG score. This is the first study investigating onco-geriatric scores regarding patient-reported outcomes. Since a high percentage of PRO was missing, all patients who could be followed at least during one cycle were considered for the analysis to include the maximum information available. However, possible misclassification of patients without complete follow-up must be considered when interpreting results. Due to the nature of patient-reported outcomes, toxicity based on laboratory variables (e.g. neutropenia) could not be measured directly via PRO. Therefore, some parts of the assessed toxicity during the development of the scores were missing. For this analysis, the patients' total severe symptom burden was considered in order to focus on the general health condition of the patient. In future studies, separate analyses of the different symptoms might be conducted for investigating if the scores can provide a higher predictive performance for one of the individual symptoms. As previously discussed in section 5.1.3, the mode of asking the questions, the time point of contacting patients, and the missing data, might limit generalizability of the results. However, since patient-reported outcomes play an increasing role in oncology, this exploratory analysis is an essential contribution to further research in this field.

Predictive performance of the onco-geriatric scores for alterations of the planned treatment

Both scores exhibited a good predictive performance for major alterations of the planned treatment and the time to those alterations. This is the first study evaluating the potential of onco-geriatric scores for the prediction of the alterations of planned treatment. Presumably, the prediction of major alterations with a score for toxicity prediction should be possible since discontinuations and changes are often caused by severe toxicity. This finding is consistent with literature where different domains of the geriatric assessment (e.g. ECOG performance status, renal function) have shown to be associated with the completion of chemotherapy as planned [133]. Interestingly, for minor alterations, the CARG score and the CRASH score demonstrated a tendency to inverse predictions: With higher risk categories, patients

experienced fewer minor alterations. This might be explained by the fact that for patients with a higher risk, and thus a worse health status, consequences of toxicity might be rather severe, leading to less minor modifications but instead rather to discontinuations or changes of therapy regimen (major modifications).

5.1.6 Prediction of severe toxicity by other predictive factors

ECOG, age, and physicians' judgement

Neither physicians' judgment nor the commonly used predictors ECOG performance status and age indicated adequate predictive performance for overall toxicity.

Two previous studies investigated the physicians' judgment vs the CARG score prediction [72, 73]. Consistent with our results, those studies did not observe an adequate predictive performance of physicians' judgment. However, both studies did not find an adequate predictive value of the CARG score either [72, 73]. Contrary to these results, in our patient cohort, the CARG score and the CRASH score predicted overall toxicity better than physicians' judgment. This result is also in line with the finding that a CGA may be more effective in selecting older patients for aggressive chemotherapy than physicians [45]. Moreover, a CGA was found to identify different patients as fit for chemotherapy than clinical judgment [143]. In general, these results underline the value of onco-geriatric scores in supporting physicians during the decision-making process. However, the results of the evaluation study also need to be differentiated between the different toxicity types: In our study, the physicians' judgment for hematologic toxicity performed better than the physicians' judgment for overall and nonhematologic toxicity.

Age and ECOG are commonly used for estimating the risk of cancer therapies in older patients [45]. In this evaluation study, age alone did not adequately predict toxicity. This is in line with our expectations since aging is a highly individualized process [13]. The ROC curve for age was located below the line of identity, suggesting that higher age was associated with less toxicity. This might be caused by the tendency to treat patient of higher age with less toxic regimens [12]. The ECOG performance status did not adequately predict toxicity neither for overall nor for hematologic toxicity. However, for nonhematologic toxicity, ECOG might be more useful.

Those results are consistent with the poor results obtained in the comparison of the CARG score with the Karnofsky performance status (KPS), another commonly used performance status. In the development study of the CARG score, the proportion of patients with overall toxicity increased for the CARG score categories (low: 30%, mid: 52%, high: 83%) whereas the KPS did not differentiate well (KPS 90-100%: 51%, KPS 80%: 51%, KPS ≤ 70%: 62% [67]). The validation study of the CARG score and a study with lung cancer patients yielded similar results regarding the KPS [71, 83]. Nevertheless, an Australian survey indicated that the performance status is still the most important factor for oncologists in decision-making [144].

Individual CARG and CRASH score items

Surprisingly, most individual score items did not indicate a significant association with severe overall toxicity when analyzed separately. This suggests a substantial synergistic power of the items since combined within one score, toxicity was mostly predicted significantly. For the CARG score, hemoglobin was the only significant item and for the CRASH score, LDH. These results are interesting since merely laboratory variables exhibited an association – although geriatric assessment variables are deemed highly important for determining the health status of an older patient [13]. The hemoglobin OR of 4.036 (1.110-14.683) was higher in our study than during the CARG development study, but the OR was located within the confidence interval of the CARG study (OR 2.31, CI 1.15-4.64) [67]. The OR of the LDH was 6.980 (1.432-34.035), being higher than the OR reported in the CRASH development study for hematologic toxicity (no OR reported for overall toxicity) [68]. However, this value might be interpreted with caution since the number of patients in the low LDH category solely consisted of seven patients in our study.

Tumor type was not associated with severe overall toxicity in this evaluation study. In contrast, Hurria et al. included tumor type (genitourinary/gastrointestinal vs others) into the CARG prediction model [67]. However, our results are in line with those by Extermann et al. who observed that tumor entities do not exhibit a substantial influence on toxicity [68]. This supports the assumption that both scores are valid across tumor entities, thus also for hematologic malignancies. The type of treatment denoted a significant association with toxicity, confirming the previous expectation that targeted or immunotherapies imply less toxicity than chemotherapy. The toxicity of treatment is already considered in the CRASH

score by the MAX2 index [68] and in the CARG score by the variable “monotherapy vs polytherapy” [67].

In literature, heterogeneous results were found regarding predictors for toxicity: A systematic review reported that in five of 13 studies, tumor type and chemotherapy regimen were associated with chemotherapy toxicity, older age was associated with toxicity in three out of nine studies, and cognitive impairment was associated with chemotherapy toxicity in two out of six studies [145]. The only significant predictors of the CARG and CRASH score in this study, hemoglobin and LDH, were not mentioned as predictors in this review. A review by Versteeg et al. did not find any consistency in predictive factors for toxicity from a geriatric assessment either [54]. Contrary to the results of our study, a secondary analysis of the CARG development study found that renal function was associated with an increase in chemotherapy-related toxicity [146].

Since some items only consisted of few patients per category, logistic regression results should be interpreted with caution. For further analysis, a study with a larger patient cohort should be conducted where a multivariate logistic regression analysis could be performed to adjust for covariates.

5.2 Medication risk analysis

Medication risks were common in older cancer patients even before start of cancer therapy: 52.2% of patients were exposed to polymedication, 52.9% to potentially inadequate medication (PIM), and 30.9% to relevant potential drug-drug interactions (rPDDI). Their prevalence increased after start of cancer therapy. rPDDI were significantly associated with the adverse outcome of severe overall and hematologic toxicity.

5.2.1 Study set-up

Patients from the pilot and evaluation study were pooled in order to increase the sample size and enhance validity of the findings. Since eligibility criteria were similar, pooling was justifiable. The only difference in eligibility criteria, the performance of tumor therapy, was considered by only including patients actually starting therapy. Since merely 16/136 (11.8%)

patients originated from the pilot study, the patient characteristics mainly corresponded to those of the evaluation study, being previously discussed in section 5.1.1.

A strength of this study was the investigation of two distinct time points which allowed analyzing the changes in medication before and after the start of cancer therapy. Moreover, we considered the association with adverse outcomes for patients, being essential for assessing the clinical implications of our findings.

Regarding limitations, the retrospective character of this analysis should be mentioned. The documentation of drugs in the medical records might have been incomplete, leading to an underestimation of drug use. Probably, more drugs – in particular nonprescription drugs (only being 96/683 in this analysis) – could have been detected if patients had been specifically interviewed by a pharmacist concerning medication use. Moreover, information on the duration and rationale of drug use was partly missing due to the retrospective design, limiting judgment of PIM. However, this concern was addressed by selecting an explicit PIM list, requiring little additional data. A further limitation is the moderate sample size of the analysis. Nevertheless, by pooling of data, a higher sample size was reached. Further studies with a prospective design and a larger patient cohort are needed to corroborate results.

5.2.2 Polymedication

On average, patients took 5 (SD, 3.5) drugs as long-term medication before start of cancer therapy. Interestingly, the cohort of the CARG score development study (n = 500) was also investigated regarding polymedication and PIM use. This secondary analysis indicated a mean of 5 (SD, 4) drugs per patient [67, 147], consistent with our results. Turner et al. investigated the medication of 385 older cancer patients via a self-reported medication data instrument which patients completed before the initial appointment at an outpatient oncology clinic. This study found a mean of 5.7 (SD, 3.7) drugs per patient, in line with our findings [25]. Nightingale et al. detected an average of 9.2 drugs per patient [38]. The higher prevalence might be explained by the different study designs: Nightingale et al. evaluated the medication being documented during a pharmacist-led comprehensive medication assessment. Since patients were advised to bring their complete medication during this session, it seems plausible that the study found a higher number of drugs. The study detected that 26.5% of patients used

complementary and alternative medication like herbal medicines or dietary supplements in this cohort [148], and on average three OTC or herbal drugs were taken [38]. Neglecting those three nonprescription drugs per patient, which our study might not have found due to the study design, Nightingale et al. presented similar results as our study.

Half of the patients demonstrated polymedication (≥ 5 drugs) and 10.3% of patients hyperpolymedication (≥ 10 drugs) in our analysis. Turner et al. found a prevalence of 57% for polymedication and 15% for hyperpolymedication [25], which is in line with our findings. Prithviraj et al. detected that 80% of older cancer patients used five or more drugs prior to start of treatment [26] whereas Alkan et al. found a prevalence of 30.8% taking at least five drugs [149]. A review by Sharma et al. indicated that polymedication prevalence for older cancer patients ranges from 11% to 96% in literature [22]. Varying definitions of polymedication might play a role for this variability [22] but also differences in data collection (e.g. interview by a pharmacist vs self-reported data collection form) and in counting of drugs (e.g. counting per active ingredient vs per medicinal product).

The two most prevalent drug classes of this study cohort were antithrombotic agents and agents acting on the renin-angiotensin system. This corresponds with the two most frequent drug classes observed by Turner et al. [25].

The high prevalence of polymedication in older cancer patients is important for health care providers to be aware of, due to its association with various risks, like frailty and decreased physical function [25, 27]. Moreover, the number of drugs was associated with a higher risk of therapy-related toxicity in some studies [22], see section 5.2.6. However, when considering the clinical impact of polymedication, it is essential to point out that a high number of drugs is not per se inappropriate for older patients [150].

5.2.3 Potentially inadequate medication

The appropriateness of medication use is considered in this study via PIM screening. Patients took in median 1 (range, 0-5) PIM drug in their long-term medication and more than half of the patients used at least one PIM drug in this analysis, according to the EU(7)-PIM list. This result is within the range of the previously reported PIM prevalence for the general older population: A systematic review found that using administrative data, the prevalence of PIM

ranged between 11.5% and 62.5%, varying widely due to different methods and cohorts [151]. Regarding older cancer patients, a secondary analysis of the CARG study cohort found that 29% of patients used at least one PIM drug (2012 Beers criteria) before start of treatment [147]. Other studies in older cancer patients detected prevalence rates of 40% and 26.5% using the 2012 Beers criteria [38, 152]; a prevalence of 38% was found following STOPP criteria [38]. Recent studies used the 2015 Beers criteria and detected a higher prevalence of PIM drugs for older cancer patients, being closer to the results of this study: Analyzing an epidemiologic database, Feng et al. detected a PIM prevalence of 61.7% for breast cancer, 47.3% for prostate cancer, and 66.3% for colorectal cancer patients during therapy [153]; Moreira Reis et al. found that 48.1% of older cancer patients receiving parenteral cancer therapy used at least one PIM drug [37]. Patients in these studies had already started cancer therapy which may have increased PIM burden. However, the 2015 Beers update also generally presents a higher PIM prevalence compared to its 2012 version since the 2015 update includes proton-pump inhibitors [154].

The EU(7)-PIM list has been used in other cohorts like cognitive impaired patients [155], but this is the first study assessing PIM use of older cancer patients with the EU(7)-PIM list. This limits comparison with literature since different explicit PIM lists are heterogeneous in content [156]. However, Morin et al. found that different explicit PIM lists may yield similar results of PIM prevalence if used at the population level [157].

The most frequently applied PIM drug class in the study by Moreira Reis et al. comprised proton-pump inhibitors (PPI) in 33.3% of patients [37], corresponding to our findings where PPI were the most prevalent PIM drugs by far. PPI are deemed inappropriate in long-term use, since they can increase the risk of Clostridium difficile infections and hip fractures [33]. The EU(7)-PIM list only classifies PPI as PIM if taken longer than 8 weeks. Mostly, information on the duration of usage was missing in this study. Since PPI are often used for a longer time period in general, it was assumed that counting them as PIM in case of missing information would reflect reality best. However, this approach might have led to an overestimation of PIM prevalence.

An inherent limitation of explicit PIM lists is the generalization of drugs being inappropriate, instead of assessing the clinical situation of an individual patient. However, due to lacking data (e.g. diagnoses), implicit PIM lists would not have been reliable in this study. Moreover,

explicit tools are more widely used in clinical routine [158], allowing for a better transportability of results into clinical practice. Nevertheless, explicit criteria may not consider new drugs being released after compiling of the PIM list. For instance the EU(7)-PIM list does not mention edoxaban although rivaroxaban or apixaban are listed [33].

Whitman et al. recommended using several PIM lists concomitantly for assessing PIM use of older patients since that bears complementary effects [125]. In a study with a pharmacist-led medication assessment, analysis via three combined tools (Beers, START/STOPP, and MAI score) discovered three times more PIM drugs than by using Beers criteria alone [159]. Future prospective studies with more clinical data, might combine the EU (7)-PIM list with other instruments like for instance the START/STOPP criteria or the FORTA list [34, 35].

The clinical impact of a high PIM burden is controversial for the cohort of older cancer patients. Two recent studies by Karuturi et al. questioned an association of PIM with adverse outcome, analyzing an epidemiologic database regarding older patients with breast and colorectal cancer. Instead, the authors hypothesized an association of outcomes with polymedication [160, 161]. However, these findings still require prospective verification. Sharma et al. acknowledged reasons for the lack of an association in geriatric oncology: For example, some PIM might not be harmful in older cancer patients but instead necessary for supportive therapy and end-of-life medicine. There also might still exist too few and too heterogeneous studies to show an association [22].

5.2.4 Relevant potential drug-drug-interactions

Approximately one third (30.9%) of patients experienced rPDDI in their long-term medication in this study. In literature, the results vary: Popa et al. observed a prevalence of 75.4% for potential drug-drug interactions of older cancer patients during therapy [43], a substantially higher value than in our analysis. This might be due to the inclusion of antineoplastic agents but probably primarily results from the inclusion of all interactions, disregarding severity grades. Of note, a large percentage of the interactions found by Popa et al. demonstrated only minor clinical significance [43]. Considering all severity grades in our study, 56.6% of patients would have shown interactions. Yeoh et al. found 55.1% of older cancer patients being exposed to potential drug-drug interactions [15]. This finding was higher than in our study and

might be explained by the fact that the study also included cancer therapy agents. Also, the study only enrolled patients taking at least three long-term drugs. In general, other studies in literature bear limited comparability due to the different methods for interaction detection. Our study used the ABDA database classification system which is very common in Germany but rather unknown in other countries. Different interaction information systems have presented deviant listing of interactions and variant severity classifications [162, 163].

Most rPDDI detected in this analysis were classified as “monitoring/modification needed” according to the ABDA database classification; no contraindications were found in the long-term medication. These are positive findings, indicating no highly serious issues for patients.

Yeoh et al. found statins and sulfonylureas to be frequent interaction partners [15]. In our study, those drug classes were also involved in frequent interactions types (anti-diabetic drugs – corticosteroids, simvastatin – amlodipine). The most frequent interaction partners in our study comprised agents acting on the renin-angiotensin system and beta blockers. These drug classes also belonged to the most frequently administered ones in general. However, health care providers should be especially vigilant about the drug classes “cardiac therapy” (ATC code C01) and “corticosteroids for systemic use” (ATC code H02), showing the highest potential of provoking interactions when prevalence was considered. The high interaction propensity of the drug class “cardiac therapy” was caused by the interaction between amiodarone and beta blockers (ABDA database category: Monitoring or modification needed) which may lead to additive cardiodepressant effects [126]. Corticosteroids indicated a high interaction propensity due to a frequent interaction between corticosteroids and antidiabetic drugs (ABDA database category: Monitoring or modification needed) [126]. The blood glucose lowering effects of antidiabetics are attenuated by corticosteroids [126].

The clinical consequences of those drug-drug interactions might be severe: A retrospective study indicated that about 2% of unplanned hospitalizations of cancer patients were caused by drug-drug interactions [44]. However, when interpreting the results of our study, it is important to consider that all these drug-drug interactions are merely potential. If they are clinically relevant remains unclear, indicating an interesting field of further research.

5.2.5 Antineoplastic agents and supportive care medication

At start of cancer therapy, patients received in median six additional drugs due to either antineoplastic agents or supportive care medication. This finding depends largely on the tumor entities and the applied therapy regimens. That limits the generalizability to other patient cohorts with e.g. mostly targeted or immunotherapy regimens, frequently comprising only a low number of antineoplastic agents.

Potentially inadequate medication

More than one third of patients received additional PIM drugs after start of therapy. The most commonly used additional PIM drug after start of cancer therapy was ranitidine, bearing the risk of CNS adverse effects like confusion [33]. If Ranitidine was only used as premedication (e.g. in paclitaxel regimens) the drug was not considered as PIM since this study only took into account supportive therapy which was used more than once per cycle for enhancing clinical relevance of findings. Nevertheless, Ranitidine still remained the most common PIM drug. Whereas this study was interested in the *additional* PIM burden by antineoplastic agents and supportive care medication, several studies in literature analyzed the overall PIM burden after start of therapy. Considering our results, it would be likely for overall PIM burden to increase after start of therapy. Interestingly, epidemiologic studies in literature do not support this expectation: Karuturi et al. found a decrease of PIM prevalence in older patients after the diagnosis of breast or colorectal cancer (PIM prevalence breast cancer: pre-chemotherapy 36.6% vs 0-3 months after start of chemotherapy 27.9% vs 3-6 months after start of chemotherapy 20%) [160]. Hence, some PIM drugs in long-term medication might be discontinued after the start of cancer therapy, reducing the overall PIM burden. For breast cancer patients, Lund et al. explained the decrease in overall PIM burden by the discontinuation of estrogen [164].

This study conducted a cross-sectional analysis at the first therapy cycle. Thus, this analysis might have underestimated the total risk of PIM arising during therapy course. However, Leger et al. observed the PIM use in 122 older patients with hematologic malignancies at start of therapy and after three months, not finding a significant difference between percentages of patients with PIM [165].

Since some PIM are required as pre-medication or supportive care medication in cancer therapy, the benefit-risk assessment of some PIM drugs in cancer patients may differ from other older patients. The NCCN guideline lists various drugs which might be of concern for elderly patients but which are commonly used for supportive therapy [19]. This includes for example corticosteroids used in the prevention of nausea and emesis [19]. About 45% of the NCCN templates for therapy regimens in hematologic cancer patients comprise at least one PIM drug [166]. Our analysis included all PIM drugs because, regardless of its use in supportive therapy, they bear certain risks in older patients which physicians should be aware of. Also, the aim of this study was to measure the risk potential of the medication, not the appropriateness of physicians' prescriptions. In literature, this issue was handled differently: Maggiore et al. separately analyzed PIM including and excluding the agents being used in supportive therapy regimens [147], Moreira Reis et al. incorporated all PIM drugs for the same reasons as in our study [37]. Only a slight difference in prevalence was found by Feng et al. when neglecting the appropriate PIM drugs for cancer patients compared to including all PIM drugs [153]. Therefore, the exclusion of those PIM drugs being commonly used for supportive therapy would probably not have altered the prevalence of PIM to a large extent. However, when interpreting the results, it should be kept in mind that in a cohort of cancer patients for some PIM drugs , the benefit-risk assessment might be positive. Maggiore et al. suggested the development of specific geriatric oncology-centric definitions of polymedication and PIM use in order to satisfy the special needs of this cohort [147].

Relevant potential drug-drug interactions

About one third of patients demonstrated additional rPDDI between the antineoplastic/supportive care medication and the long-term medication after start of cancer therapy. Those additional rPDDI included higher severity grades than the rPDDI between only long-term medication. Therefore, it is important to review medication for new rPDDI after initiating cancer therapy. The results of this analysis advise particular caution when prescribing serotonin 5-HT3 receptor antagonists due to the severity of triggered rPDDI. The two most severe interaction types were both caused by serotonin 5-HT3 receptor antagonists due to their QT prolonging properties (QT prolonging drug – antiarrhythmic agent: Serious consequences possible – as precaution contraindicated (n = 3); QT prolonging drug – antidepressant: Simultaneous usage not recommended (n = 3)). The QT prolongation might

be of special concern in older patients who commonly demonstrate cardiovascular risk factors. Moreover, special attention is required when administering corticosteroids and antibiotics. Whereas corticosteroids were the drug class most frequently causing rPDDI (most frequent interactions: NSAID – corticosteroids; anti-diabetic drugs – corticosteroids), antibiotics were the drug class with the highest interaction propensity due to numerous interactions between trimethoprim and ACE inhibitors/sartans causing a higher risk of hyperkalemia [126].

5.2.6 Association of long-term medication with severe toxicity

Occurrence of polymedication in long-term medication was not associated with either overall, hematologic, or nonhematologic grade ≥ 3 toxicity in our analysis. However, the number of drugs per patient was associated with nonhematologic toxicity. In literature, results were not consistent regarding the association between the number of drugs and severe toxicity in older cancer patients [22]. In line with our results, a secondary analysis of the CARG development study did not find an association of the number of daily drugs before start of chemotherapy and overall chemotherapy-related toxicity [147]. In contrast, Hamaker et al. detected a significant association between baseline polymedication and severe toxicity during cancer treatment of older metastatic breast cancer patients [28].

PIM use was associated with grade ≥ 3 toxicity for nonhematologic toxicity only. Similar results are found in literature: Maggiore et al. did not report any association between PIM use and overall grade ≥ 3 toxicity, regardless of the applied criteria (categorized according to Beers 2012, Zhan criteria, and Drugs to Avoid in the Elderly (DAE) criteria) [147]. Likewise, Park et al. did not find any relationship between PIM use (Beers 2012) and treatment-related toxicity in head and neck cancer patients of high age [167].

Occurrence of rPDDI was significantly associated with grade ≥ 3 overall and hematologic toxicity in our study. However, all but two of the patients with rPDDI experienced severe toxicity. This uneven distribution might have influenced results for overall toxicity. A study by Popa et al. indicated that potential drug-drug interactions were not associated with grade 4 hematologic toxicity [43]. In contrast, grade ≥ 3 nonhematologic toxicity was significantly associated with potential drug-drug interactions of higher severity (“level 1-3”) in that study.

These results might differ from our study because another software for classifying potential drug-drug interactions was used.

Because different medication risks, especially drug-drug interactions, suggest an association with severe toxicity in this cohort, interventions for improving prescribing quality are warranted. Assessing polymedication, PIM, or drug-drug interactions can reveal essential risks but does not show the whole picture regarding quality of medication use. Thus, interventions should preferably consist of a medication review going beyond simple drug counting and PIM lists, as recommended in the NCCN guidelines [19]. If possible, a multidisciplinary approach including a clinical pharmacist should be pursued [41, 168]. First results of interventions are promising: For older cancer patients, Deliens et al. found that a clinical pharmacist could significantly reduce PIM of patients hospitalized in a geriatric oncology unit [169]. Nightingale et al. reported a reduction of the average number of drug-related problems by 45.5% via a pharmacist-led, individualized medication assessment [17]. For cancer patients in general, a systematic review concluded that interventions by pharmacists may improve outcomes of cancer patients [170]. Further studies are necessary to investigate if interventions may also improve outcomes in the cohort of older cancer patients.

A limitation of this analysis is the relatively small sample size and the univariate approach. A multivariate analysis in a larger sample size should be conducted to corroborate results. Moreover, apart from toxicity, other patient-relevant endpoints like hospitalization or survival could be of interest for further analyses.

5.3 Conclusion and outlook

This is the first study directly comparing the CARG and the CRASH score for the prediction of therapy-related toxicity at advanced age in the clinical routine setting.

As both onco-geriatric scores presented a similar predictive performance, in general, none could be recommended above the other. However, the CARG score needs less time and hence might be preferable in busy routine due to its ease of use [62]. The hematologic CRASH score, in contrast, should be preferred if a more detailed estimation of hematologic toxicity is required. Despite a similar predictive performance, the onco-geriatric scores predicted

different risks, thus not being completely interchangeable. Both scores suggested better toxicity predictions than the physicians' judgment alone and hence might be useful in supporting clinical decisions. Moreover, the onco-geriatric scores indicated a better predictive performance for overall toxicity than the commonly used predictors ECOG performance status and age.

This study indicated that the CARG and the CRASH score may be applied in a patient cohort with hematologic tumors and combinations of chemotherapy with targeted or immunotherapies. Since targeted therapy and immunotherapy play an increasing role in cancer therapy, future studies should investigate the toxicity prediction for those therapies in more detail. Exploratory analyses suggested that both scores may also be useful for predicting the time until onset of severe toxicity, as well as the occurrence of major therapy alterations of the planned treatment (discontinuations or changes). Furthermore, an explorative analysis found that the CRASH score might be valuable for predicting severe patient-reported symptom burden. Since patients have shown to report severe toxicity earlier and more frequently than physicians [97], it might be interesting to further investigate the application of the onco-geriatric scores for predicting patient-reported outcomes.

The analysis of medication risks in older cancer patients indicated that even before start of cancer therapy, medication risks were common: More than half of patients were exposed to polymedication and potentially inadequate medication (PIM) use; one third exhibited relevant potential drug-drug interactions (rPDDI). The most frequent PIM drugs in long-term medication were proton-pump inhibitors. The drug classes with the highest potential of provoking interactions comprised cardiac therapy (ATC code C01) and corticosteroids for systemic use (ATC code H02). After start of cancer therapy, the risks in the medication increased: One third of patients exhibited additional PIM drugs and one third additional rPDDI. Regarding supportive therapy, especially serotonin antagonists and antibiotics should be used with caution, due to their high potential of interactions, respectively. In a univariate analysis, the occurrence of rPDDI was associated with a higher risk for patients to develop overall or hematologic toxicity. A multivariate analysis in a larger sample size should be conducted to further investigate the association between medication risks and severe toxicity in this cohort. Showing that risks in the medication of older cancer patients are common and may be

associated with toxicity, this raises the need for multi-disciplinary interventions to optimize medication use in this cohort.

To conclude, by assessing onco-geriatric scores for toxicity prediction and by evaluating medication risks, this work contributed to further improvements in the pharmacotherapy of older cancer patients.

6 Summary

The cancer therapy of older patients is challenging, being more complex than the therapy of younger patients. Older cancer patients show a higher toxicity risk during therapy and drug-related problems are common. In order to individualize cancer care in this heterogeneous population, short tools combining geriatric assessment with oncologic parameters were developed for predicting toxicity during chemotherapy: the CARG (Cancer and Aging Research Group) and the CRASH (Chemotherapy Risk Assessment Scale for High Age Patients) score. The aim was to compare the scores regarding their predictive performance in a clinical routine setting. Moreover, this thesis aimed at evaluating medication-related risks in older cancer patients.

In a prospective, single-center observational study, the CARG and the CRASH score were assessed for patients ≥ 70 years before the start of their systemic cancer treatment. The CARG score predicts severe overall toxicity. The CRASH score is divided into three subcategories, predicting severe overall, hematologic, and nonhematologic toxicity. Moreover, physicians' judgments regarding the patients' toxicity risk were documented. Grade ≥ 3 toxicity according to the Common Terminology Criteria for Adverse Events (CTCAE) was captured from medical records. The predictive performance of the scores was assessed by analyzing the proportion of patients with severe toxicity per risk category, logistic regression, and the area under the receiver operating characteristic curve (ROC-AUC). Furthermore, the prediction performance was compared with other commonly used predictors. The evaluation study of the CARG and the CRASH score comprised 120 patients (50% female, mean age 77.2 years). Severe toxicity was experienced by 81% of patients; 67% showed signs of hematologic toxicity. The predictive performances of the CARG score and the combined CRASH score were similar for overall and nonhematologic toxicity. For hematologic toxicity, the hematologic CRASH score performed better than the CARG score. Neither physicians' judgment nor the ECOG nor age indicated adequate predictive performance for overall toxicity.

Medication risks in older cancer patients were investigated regarding polymedication (defined as the use of ≥ 5 drugs), potentially inadequate medication (PIM; defined by the EU(7)-PIM list), and relevant potential drug-drug interactions (rPDDI; analyzed by the ABDA interaction database). Before the start of cancer therapy, patients took on average 5 drugs as long-term medication and 52% of patients were exposed to polymedication. More than half of patients

used at least one PIM (mostly drugs for acid-related disorders). Approximately one third of patients experienced rPDDI.

In conclusion, the CARG and the CRASH score exhibited similar predictive performance for overall and nonhematologic toxicity. However, the hematologic CRASH score should be preferred for predicting hematologic toxicity. Both scores performed better than clinical judgment alone and thus may be used for supporting therapy decisions in clinical routine. Medication risks were common in older cancer patients, raising the need for interdisciplinary interventions to ensure medication safety in this cohort.

7 References

1. Robert-Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V., editors. Krebs in Deutschland für 2013/2014. 11rd ed. Berlin; 2017 [cited 2019 Sep 28]. Available from: http://www.krebsdaten.de/.

2. Ebert M, Härtel N, Wedding U, editors. Geriatrische Onkologie. 1st ed. Berlin, Heidelberg: Springer-Verlag; 2018.

3. Nightingale G, Schwartz R, Kachur E, Dixon BN, Cote C, Barlow A et al. Clinical pharmacology of oncology agents in older adults: a comprehensive review of how chronologic and functional age can influence treatment-related effects. J Geriatr Oncol 2019; 10:4–30.

4. Wedding U, Roehrig B, Klippstein A, Steiner P, Schaeffer T, Pientka L et al. Comorbidity in patients with cancer: Prevalence and severity measured by cumulative illness rating scale. Crit Rev Oncol Hematol 2007; 61:269–76.

5. Williams GR, Mackenzie A, Magnuson A, Olin R, Chapman A, Mohile S et al. Comorbidity in older adults with cancer. J Geriatr Oncol 2016; 7:249–57.

6. Chaïbi P, Magné N, Breton S, Chebib A, Watson S, Duron JJ et al. Influence of geriatric consultation with comprehensive geriatric assessment on final therapeutic decision in elderly cancer patients. Crit Rev Oncol Hematol 2011; 79:302–7.

7. Scher KS, Hurria A. Under-representation of older adults in cancer registration trials: Known problem, little progress. J Clin Oncol 2012; 30:2036–8.

8. Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ et al. Participation of patients 65 years of age or older in cancer clinical trials. J Clin Oncol 2003; 21:1383–9.

9. Nipp RD, Yao N, Lowenstein LM, Buckner JC, Parker IR, Gajra A et al. Pragmatic study designs for older adults with cancer: report from the U13 conference. J Geriatr Oncol 2016; 7:234–41.

10. Lichtman SM, Hurria A, Cirrincione CT, Seidman AD, Winer E, Hudis C et al. Paclitaxel

efficacy and toxicity in older women with metastatic breast cancer: combined analysis of CALGB 9342 and 9840. Ann Oncol 2012; 23:632–8.

11. Muss HB, Berry DA, Cirrincione C, Budman DR, Henderson IC, Citron ML et al. Toxicity of older and younger patients treated with adjuvant chemotherapy for node-positive breast cancer: the cancer and leukemia group B experience. J Clin Oncol 2007; 25:3699–704.

12. Bouchardy C, Rapiti E, Fioretta G, Laissue P, Neyroud-Caspar I, Schäfer P et al. Undertreatment strongly decreases prognosis of breast cancer in elderly women. J Clin Oncol 2003; 21:3580–7.

13. Extermann M, Aapro M, Bernabei R, Cohen HJ, Droz JP, Lichtman S et al. Use of comprehensive geriatric assessment in older cancer patients: recommendations from the task force on CGA of the International Society of Geriatric Oncology (SIOG). Crit Rev Oncol Hematol 2005; 55:241–52.

14. Quaglia A, Tavilla A, Shack L, Brenner H, Janssen-Heijnen M, Allemani C et al. The cancer survival gap between elderly and middle-aged patients in Europe is widening. Eur J Cancer 2009; 45:1006–16.

15. Yeoh TT, Tay XY, Si P, Chew L. Drug-related problems in elderly patients with cancer receiving outpatient chemotherapy. J Geriatr Oncol 2015; 6:280–7.

16. Pharmaceutical Care Network Europe Association. Classification for drug related Problems. V.8.02. [cited 2019 May 19]. Available from: http://www.pcne.org/.

17. Nightingale G, Hajjar E, Pizzi LT, Wang M, Pigott E, Doherty S et al. Implementing a pharmacist-led, individualized medication assessment and planning (iMAP) intervention to reduce medication related problems among older adults with cancer. J Geriatr Oncol 2017; 8:296–302.

18. Buajordet I, Ebbesen J, Erikssen J, Brørs O, Hilberg T. Fatal adverse drug events: the paradox of drug treatment. J Intern Med 2001; 250:327–41.

19. NCCN practice guidelines in oncology: Older Adult Oncology, version 1.2019 [cited 2019 Mar 2]. Available from: http://www.nccn.org/.

20. Wildiers H, Heeren P, Puts M, Topinkova E, Janssen-Heijnen MLG, Extermann M et al. International Society of Geriatric Oncology consensus on geriatric assessment in older patients with cancer. J Clin Oncol 2014; 32:2595–603.

21. Hohmann C. Geriatrische Pharmazie. In: Jaehde U, Radziwill R, Kloft C, editors. Klinische Pharmazie. 4th ed. Stuttgart: Wissenschaftliche Verlagsgesellschaft Stuttgart; 2017. p. 402–18.

22. Sharma M, Loh KP, Nightingale G, Mohile SG, Holmes HM. Polypharmacy and potentially inappropriate medication use in geriatric oncology. J Geriatr Oncol 2016; 7:346–53.

23. Schwabe U, Paffrath D, editors. Arzneiverordnungs-Report 2016. Berlin, Heidelberg: Springer-Verlag; 2016.

24. Thürmann P, Selke G. Arzneimittelversorgung älterer Patienten. In: Klauber J, Günster C, Gerste B, Robra B, Schmacke N, editors. Versorgungs-Report 2013/2014. Stuttgart: Schattauer; 2014. p. 185–208.

25. Turner JP, Shakib S, Singhal N, Hogan-Doran J, Prowse R, Johns S et al. Prevalence and factors associated with polypharmacy in older people with cancer. Support Care Cancer 2014; 22:1727–34.

26. Prithviraj GK, Koroukian S, Margevicius S, Berger NA, Bagai R, Owusu C. Patient characteristics associated with polypharmacy and inappropriate prescribing of medications among older adults with cancer. J Geriatr Oncol 2012; 3:228–37.

27. Hersh LR, Beldowski K, Hajjar ER. Polypharmacy in the geriatric oncology population. Curr Oncol Rep 2017; 19:73.

28. Hamaker ME, Seynaeve C, Wymenga ANM, Van Tinteren H, Nortier JWR, Maartense E et al. Baseline comprehensive geriatric assessment is associated with toxicity and survival in elderly metastatic breast cancer patients receiving single-agent chemotherapy: Results from the OMEGA study of the Dutch Breast Cancer Trialists' Group. Breast 2014; 23:81–7.

29. Cadogan CA, Ryan C, Hughes CM. Appropriate polypharmacy and medicine safety:

when many is not too many. Drug Saf 2016; 39:109–16.

30. Kaufmann CP, Tremp R, Hersberger KE, Lampert ML. Inappropriate prescribing: a systematic overview of published assessment tools. Eur J Clin Pharmacol 2014; 70:1–11.

31. American Geriatrics Society 2015 Beers Criteria Update Expert Panel. American Geriatrics Society 2015 updated beers criteria for potentially inappropriate medication use in older adults. J Am Geriatr Soc 2015; 63:2227–46.

32. Holt S, Schmiedl S, Thürmann PA. Potentially inappropriate medications in the elderly: The PRISCUS list. Dtsch Arztebl Int 2010; 107:543–51.

33. Renom-Guiteras A, Meyer G, Thürmann PA. The EU(7)-PIM list: A list of potentially inappropriate medications for older people consented by experts from seven European countries. Eur J Clin Pharmacol 2015; 71:861–75.

34. Wehling M. No Drug therapy in the elderly: too much or too little, what to do? A new assessment system: fit for the aged FORTA. Dtsch Med Wochenschr 2008; 133:2289–2291.

35. Gallagher P, Ryan C, Byrne S, Kennedy J, O'Mahony D. STOPP (Screening Tool of Older Person's Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment). Consensus validation. Int J Clin Pharmacol Ther 2008; 46:72–83.

36. Hanlon J, Schmader K, Samsa G, Weinberger M, Uttech K, Lewis I et al. A method for assessing drug therapy appropriateness. J Clin Epidemiol 1992; 45:1045–1051.

37. Moreira Reis C, Gouvêa dos Santos A, de Jesus Souza P, Moreira Reis AM. Factors associated with the use of potentially inappropriate medications by older adults with cancer. J Geriatr Oncol 2017; 8:303–7.

38. Nightingale G, Hajjar E, Swartz K, Andrel-Sendecki J, Chapman A. Evaluation of a pharmacist-led medication assessment used to identify prevalence of and associations with polypharmacy and potentially inappropriate medication use among ambulatory senior adults with cancer. J Clin Oncol 2015; 33:1453–9.

39. Reich O, Rosemann T, Rapold R, Blozik E, Senn O. Potentially inappropriate medication

use in older patients in swiss managed care plans: Prevalence, determinants and association with hospitalization. PLoS One 2014; 9(8):e105425.

40. Hamilton H, Gallagher P, Ryan C, Bryne S, O'mahony D. Potentially inappropriate medications defined by STOPP Criteria and the risk of adverse drug events in older hospitalized patients. Arch Intern Med 2011; 171:1013–9.

41. Lees J, Chan A. Polypharmacy in elderly patients with cancer: clinical implications and management. Lancet Oncol 2011; 12:1249–57.

42. Riechelmann RP, Tannock IF, Wang L, Saad ED, Taback NA, Krzyzanowska MK. Potential drug interactions and duplicate prescriptions among cancer patients. J Natl Cancer Inst 2007; 99:592–600.

43. Popa MA, Wallace KJ, Brunello A, Extermann M, Balducci L. Potential drug interactions and chemotoxicity in older patients with cancer receiving chemotherapy. J Geriatr Oncol 2014; 5:307–14.

44. Miranda V, Fede A, Nobuo M, Ayres V, Giglio A, Miranda M et al. Adverse drug reactions and drug interactions as causes of hospital admission in oncology. J Pain Symptom Manage 2011; 42:342–53.

45. Tucci A, Ferrari S, Bottelli C, Borlenghi E, Drera M, Rossi G. A comprehensive geriatric assessment is more effective than clinical judgment to identify elderly diffuse large cell lymphoma patients who benefit from aggressive therapy. Cancer 2009; 115:4547–53.

46. Kim J, Hurria A. Determining chemotherapy tolerance in older patients with cancer. J Natl Compr Canc Netw 2013; 11:1494–502.

47. Jolly TA, Deal AM, Nyrop KA, Williams GR, Pergolotti M, Wood WA et al. Geriatric assessment-identified deficits in older cancer patients with normal performance status. Oncologist 2015; 20:379–85.

48. Puts MTE, Alibhai SMH. Fighting back against the dilution of the comprehensive geriatric assessment. J Geriatr Oncol 2018; 9:3–5.

49. Zubrod CG, Schneiderman M, Frei E, Brindley C, Gold GL, Shnider B et al. Appraisal of methods for the study of chemotherapy of cancer in man: Comparative therapeutic trial

of nitrogen mustard and triethylene thiophosphoramide. J Chron Dis 1960; 11:7–33.

50. Karnofsky D, Burchenal J. The clinical evaluation of chemotherapeutic agents in cancer. In: Macleod C, editor. Evaluation of chemotherapeutic agents. New York: Columbia University Press; 1949. p. 191–205.

51. ECOG-ACRIN. Comparing the ECOG Performance Status to the Karnofsky Performance Status [cited 2019 Sep 29]. Available from: https://www.ecog-acrin.org/.

52. Ellis G, Whitehead MA, O'Neill D, Langhorne P, Robinson D. Comprehensive geriatric assessment for older adults admitted to hospital. Cochrane Database Syst Rev 2011; (7):1–86.

53. Kenis C, Bron D, Libert Y, Decoster L, Van Puyvelde K, Scalliet P et al. Relevance of a systematic geriatric screening and assessment in older patients with cancer: results of a prospective multicentric study. Ann Oncol 2013; 24:1306–12.

54. Versteeg KS, Konings IR, Lagaay AM, van de Loosdrecht AA, Verheul HMW. Prediction of treatment-related toxicity and outcome with geriatric assessment in elderly patients with solid malignancies treated with chemotherapy: a systematic review. Ann Oncol 2014; 25:1914–8.

55. Hamaker ME, te Molder M, Thielen N, van Munster BC, Schiphorst AH, van Huis LH. The effect of a geriatric evaluation on treatment decisions and outcome for older cancer patients – A systematic review. J Geriatr Oncol 2018; 9:430–40.

56. Caillet P, Laurent M, Bastuji-Garin S, Liuu E, Culine S, Lagrange J-L et al. Optimal management of elderly cancer patients: usefulness of the comprehensive geriatric assessment. Clin Interv Aging 2014; 9:1645–60.

57. Hamaker ME, Vos AG, Smorenburg CH, de Rooij SE, van Munster BC. The value of geriatric assessments in predicting treatment tolerance and all-cause mortality in older patients with cancer. Oncologist 2012; 17:1439–49.

58. Corre R, Greillier L, Le Caer H, Audigier-Valette C, Baize N, Bérard H et al. Use of a comprehensive geriatric assessment for the management of elderly patients with advanced non-small-cell lung cancer: the Phase III randomized ESOGIA-GFPC-GECP 08-

02 study. J Clin Oncol 2016; 34:1476–83.

59. Kalsi T, Babic-Illman G, Ross PJ, Maisey NR, Hughes S, Fields P et al. The impact of comprehensive geriatric assessment interventions on tolerance to chemotherapy in older people. Br J Cancer 2015; 112:1435–44.

60. Magnuson A, Lemelman T, Pandya C, Goodman M, Noel M, Tejani M et al. Geriatric assessment with management intervention in older adults with cancer: a randomized pilot study. Support Care Cancer 2018; 26: 605–13.

61. Hamaker ME, Prins MC, Stauder R. The relevance of a geriatric assessment for elderly patients with a haematological malignancy - a systematic review. Leuk Res 2014; 38:275–83.

62. Almodovar T, Teixeira E, Barroso A, Soares M, Queiroga HJ, Cavaco-Silva J et al. Elderly patients with advanced NSCLC: The value of geriatric evaluation and the feasibility of CGA alternatives in predicting chemotherapy toxicity. Pulmonology 2019; 25:40–50.

63. Schmidt H, Boese S, Lampe K, Jordan K, Fiedler E, Müller-Werdan U et al. Trans sectoral care of geriatric cancer patients based on comprehensive geriatric assessment and patient-reported quality of life - results of a multicenter study to develop and pilot test a patient-centered interdisciplinary care concept for geriatric oncology patients (PIVOG). J Geriatr Oncol 2017; 8:262–70.

64. Hamaker ME, Jonker JM, de Rooij SE, Vos AG, Smorenburg CH, van Munster BC. Frailty screening methods for predicting outcome of a comprehensive geriatric assessment in elderly patients with cancer: a systematic review. Lancet Oncol 2012; 13:e437–44.

65. Decoster L, Van Puyvelde K, Mohile S, Wedding U, Basso U, Colloca G et al. Screening tools for multidimensional health problems warranting a geriatric assessment in older cancer patients: an update on SIOG recommendations. Ann Oncol 2015; 26:288–300.

66. Puts MTE, Tapscott B, Fitch M, Howell D, Monette J, Wan-Chow-Wah D et al. A systematic review of factors influencing older adults' decision to accept or decline cancer treatment. Cancer Treat Rev 2015; 41:197–215.

67. Hurria A, Togawa K, Mohile SG, Owusu C, Klepin HD, Gross CP et al. Predicting

chemotherapy toxicity in older adults with cancer: A prospective multicenter study. J Clin Oncol 2011; 29:3457–65.

68. Extermann M, Boler I, Reich RR, Lyman GH, Brown RH, DeFelice J et al. Predicting the risk of chemotherapy toxicity in older patients: The chemotherapy risk assessment scale for high-age patients (CRASH) score. Cancer 2012; 118:3377–86.

69. Extermann M, Bonetti M, Sledge GW, O'Dwyer PJ, Bonomi P, Benson AB. MAX2 - a convenient index to estimate the average per patient risk for chemotherapy toxicity: validation in ECOG trials. Eur J Cancer 2004; 40:1193–8.

70. Mohile SG, Dale W, Somerfield MR, Schonberg MA, Boyd CM, Burhenn PS et al. Practical assessment and management of vulnerabilities in older patients receiving chemotherapy: Asco guideline for geriatric oncology. J Clin Oncol 2018; 36:2326–47.

71. Nie X, Liu D, Li Q, Bai C. Predicting chemotherapy toxicity in older adults with lung cancer. J Geriatr Oncol 2013; 4:334–9.

72. Alibhai SMH, Aziz S, Manokumar T, Timilshina N, Breunis H. A comparison of the CARG tool, the VES-13, and oncologist judgment in predicting grade 3 + toxicities in men undergoing chemotherapy for metastatic prostate cancer. J Geriatr Oncol 2017; 8:31–6.

73. Moth EB, Kiely BE, Stefanic N, Naganathan V, Martin A, Grimison P et al. Predicting chemotherapy toxicity in older adults: Comparing the predictive value of the CARG toxicity score with oncologists' estimates of toxicity based on clinical judgement. J Geriatr Oncol 2019; 10:210–5.

74. Nishijima TF, Deal AM, Williams GR, Sanoff HK, Nyrop KA, Muss HB. Chemotherapy toxicity risk score for treatment decisions in older adults with advanced solid tumors. Oncologist 2018; 23:573–9.

75. Gesetz über den Verkehr mit Arzneimitteln (AMG) §4 [cited 2019 Sep 29]. Available from: https://www.gesetze-im-internet.de/.

76. Das Johanniter-Krankenhaus [cited 2019 Sep 29]. Available from: https://www.johanniter.de/.

77. Johanniter Krankenhaus Bonn: Zertifiziertes Onkologisches Zentrum (DKG) [cited 2019 Sep 29]. Available from: https://www.johanniter.de/.

78. Theissen Mendel M. Interdisciplinary screening and assessment for establishing age-adapted therapy concepts in elderly cancer patients [Master thesis]. Bonn: Rheinische Friedrich-Wilhelms-Universität; 2015.

79. Vergouwe Y, Steyerberg EW, Eijkemans MJC, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol 2005; 58:475–83.

80. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in logitudinal studies: development and validation. J Chroic Dis 1987; 40:373–83.

81. Extermann M. Measuring comorbidity in older cancer patients. Eur J Cancer 2000; 36:453–71.

82. Extermann M, Overcash J, Lyman GH, Parr J, Balducci L. Comorbidity and functional status are independent in older cancer patients. J Clin Oncol 1998; 16:1582–7.

83. Hurria A, Mohile S, Gajra A, Klepin H, Muss H, Chapman A et al. Validation of a prediction tool for chemotherapy toxicity in older adults with cancer. J Clin Oncol 2016; 34:2366–71.

84. National Cancer Institute: Cancers by Body Location/System [cited 2019 Sep 29]. Available from: https://www.cancer.gov/.

85. Jelliffe RW. Estimation of creatinine clearance when urine cannot be collected. Lancet 1971; 297:975–6.

86. Lawton MP. Scales to measure competence in everyday activities. Psychopharmacol Bull 1988; 24:609–14.

87. Folstein MF, Folstein SE, McHugh PR. “Mini-Mental State”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12:189–98.

88. Guigoz Y. The Mini Nutritional Assessment (MNA) review of the literature - What does it tell us? J Nutr Health Aging 2006; 10:466–87.

89. Extermann, M: How to calculate the MAX2 index and the average risk of toxicity for a chemotherapy regimen [cited 2019 Sep 29]. Available from: https://moffitt.org/.

90. DGHO: Mini-Mental State Examination (MMSE) [cited 2019 Sep 29]. Available from: https://www.onkopedia.com/.

91. Nestlé Nutrition Institute: MNA (Mini Nutritional Assessment) [cited 2019 Sep 29]. Available from: https://www.mna-elderly.com/.

92. CRASH (Chemotherapy Risk Age Scale for High Risk Patients) Scoring Analysis [cited 2019 Sep 29]. Available from: https://moffitt.org/.

93. U.S. Department of Health and Human Services: Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 (v4.03 2010) [cited 2019 Sep 29]. Available from: http://ctep.cancer.gov/.

94. National Cancer Institute: The PRO-CTCAE Measurement System [cited 2019 Sep 29]. Available from: https://healthcaredelivery.cancer.gov/.

95. Extermann M, Reich RR, Sehovic M. Chemotoxicity recurrence in older patients: risk factors and effectiveness of preventive strategies-a prospective study. Cancer 2015; 121:2984–92.

96. Di Maio M, Gallo C, Leighl NB, Piccirillo MC, Daniele G, Nuzzo F et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. J Clin Oncol 2015; 33:910–5.

97. Basch E, Jia X, Heller G, Barz A, Sit L, Fruscione M et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. J Natl Cancer Inst 2009; 101:1624–32.

98. Hagelstein V, Ortland I, Wilmer A, Mitchell SA, Jaehde U. Validation of the German patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE™). Ann Oncol 2016; 27:2294–9.

99. National Cancer Institute: Item Library (Version 1.0). Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) [cited 2019 Sep 29]. Available from: https://healthcaredelivery.cancer.gov/.

100. Reeve BB, Mitchell SA, Dueck AC, Basch E, Cella D, Reilly CM et al. Recommended patient-reported core set of symptoms to measure in adult cancer treatment trials. J Natl Cancer Inst 2014; 106:dju129.

101. Fayers P, Aaronson N, Bjordal K, Groenvold M, Curran D, Bottomley A et al. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer; 2001.

102. Bennett A V, Dueck AC, Mitchell SA, Mendoza TR, Reeve BB, Atkinson TM et al. Mode equivalence and acceptability of tablet computer-, interactive voice response system-, and paper-based administration of the U.S. National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). Health Qual Life Outcomes 2016; 14(24).

103. Lange S, Bender R. Lineare Regression und Korrelation. Dtsch Med Wochenschr 2007; 132:e9–11.

104. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960; 20:37–46.

105. Grouven U, Bender R, Ziegler A, Lange S. Der Kappa-Koeffizient. Dtsch Med Wochenschr 2007; 132:e65–8.

106. Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull 1968; 70:213–20.

107. Agresti A. Categorial data analysis. 2nd ed. New York: Wiley; 2002.

108. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159–74.

109. Field A. Discovering statistics using SPSS. 3rd ed. Los Angeles, London, New Delhi, Singapore, Washington DC: SAGE Publications; 2009.

110. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. BMJ 2009; 338:b605.

111. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Med Res Methodol 2014; 14(40).

112. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. BMJ 2009; 338:b604.

113. Bender R, Ziegler A, Lange S. Logistische Regression. Dtsch Med Wochenschr 2007; 132:e33–5.

114. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143:29–36.

115. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatr 2007; 96:644–7.

116. Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery 2016; 159:1638–45.

117. Youden WJ. Index for rating diagnostic tests. Cancer 1950; 3:32–5.

118. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958; 53:457–81.

119. Ziegler A, Lange S, Bender R. Überlebenszeitanalyse: Eigenschaften und Kaplan-Meier Methode. Dtsch Med Wochenschr 2007; 132:e36–8.

120. Ziegler A, Lange S, Bender R. Überlebenszeitanalyse: Der Log-Rang-Test. Dtsch Med Wochenschr 2007; 132:e39–41.

121. Ziegler A, Lange S, Bender R. Überlebenszeitanalyse: Die Cox-Regression. Dtsch Med Wochenschr 2007; 132:e42–4.

122. Cox DR. Regression models and life-tables. J R Stat Soc Ser B Stat Methodol 1972; 34:187–220.

123. DIMDI: Amtlicher ATC-Index 2017 [cited 2018 Mar 13]. Available from: https://www.dimdi.de/.

124. Anlage I zum Abschnitt F der Arzneimittelrichtlinie (§34 Abs.1 Satz 2 SGB V; Stand 06/2013). OTC-Übersicht [cited 2018 Mar 13]. Available from: https://www.g-ba.de/.

125. Whitman AM, DeGregory KA, Morris AL, Ramsdale EE. A comprehensive look at polypharmacy and medication screening tools for the older cancer patient. Oncologist 2016; 21:723–30.

126. ABDATA Pharma-Daten-Service. ABDA-Datenbank - Interaktionen [cited 2018 Apr 9]. Available from: http://www.abdata.de/.

127. Fargeas JB, Penot A, Olivrie A, Picat M-A, Touati M, Signol N et al. CRASH score in the older French non-Hodgkin lymphoma receiving chemotherapy, first results. J Geriatr Oncol 2014; 5:S12–3.

128. Trarieux-Signol S, Fargeas JB, Abraham J, Olivrie A, Picat MA, Signol N et al. CRASH score in the older french non Hodgkin lymphoma receiving chemotherapy, feasability. J Geriatr Oncol 2013; 4:S59.

129. Atkinson TM, Ryan SJ, Bennett AV, Stover AM, Saracino RM, Rogak LJ et al. The association between clinician-based common terminology criteria for adverse events (CTCAE) and patient-reported outcomes (PRO): a systematic review. Support Care Cancer 2016; 24:3669–76.

130. Falchook AD, Green R, Knowles ME, Amdur RJ, Mendenhall W, Hayes DN et al. Comparison of patient- and practitioner-reported toxic effects associated with chemoradiotherapy for head and neck cancer. JAMA Otolaryngol Head Neck Surg 2016; 142:517–23.

131. Flores LT, Bennett A V., Law EB, Hajj C, Griffith MP, Goodman KA. Patient-reported outcomes vs. clinician symptom reporting during chemoradiation for rectal cancer. Gastrointest Cancer Res 2012; 5:119–24.

132. Moon DH, Chera BS, Deal AM, Wang Y, Muss HB, VanderWalde NA. Clinician-observed and patient-reported toxicities and their association with poor tolerance to therapy in

older patients with head and neck or lung cancer treated with curative radiotherapy. J Geriatr Oncol 2019; 10:42–7.

133. Wildes TM, Ruwe AP, Fournier C, Gao F, Carson KR, Piccirillo JF et al. Geriatric assessment is associated with completion of chemotherapy, toxicity, and survival in older adults with cancer. J Geriatr Oncol 2013; 4:227–34.

134. Won HS, Sun DS, Choi JY, An HJ, Ko YH. Factors associated with treatment interruption in elderly patients with cancer. Korean J Intern Med 2019; 34:156–64.

135. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014; 35:1925–31.

136. Stevenson JM, Williams JL, Burnham TG, Prevost AT, Schiff R, Erskine SD et al. Predicting adverse drug reactions in older adults; a systematic review of the risk prediction models. Clin Interv Aging 2014; 9:1581–93.

137. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: A review. J Clin Epidemiol 2008; 61:1085–94.

138. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009; 338:b606.

139. Brunello A, Fontana A, Zafferri V, Panza F, Fiduccia P, Basso U et al. Development of an oncological-multidimensional prognostic index (Onco-MPI) for mortality prediction in older cancer patients. J Cancer Res Clin Oncol 2016; 142:1069–77.

140. Moth EB, Kiely BE, Stefanic N, Naganathan V, Martin A, Grimison P et al. Oncologists' perceptions on the usefulness of geriatric assessment measures and the CARG toxicity score when prescribing chemotherapy for older patients with cancer. J Geriatr Oncol 2019; 10:210–5.

141. Hentschel L, Rentsch A, Lenz F, Hornemann B, Schmitt J, Baumann M et al. A questionnaire study to assess the value of the vulnerable elders survey, G8, and predictors of toxicity as screening tools for frailty and toxicity in geriatric cancer patients. Oncol Res Treat 2016; 39:210–6.

142. Mohile SG, Velarde C, Hurria A, Magnuson A, Lowenstein L, Pandya C et al. Geriatric

assessment-guided care processes for older adults: a delphi consensus of geriatric oncology experts. J Natl Compr Canc Netw 2015; 13:1120–30.

143. Wedding U, Ködding D, Pientka L, Steinmetz HT, Schmitz S. Physicians' judgement and comprehensive geriatric assessment (CGA) select different patients as fit for chemotherapy. Crit Rev Oncol Hematol 2007; 64:1–9.

144. Moth EB, Kiely BE, Naganathan V, Martin A, Blinman P. How do oncologists make decisions about chemotherapy for their older patients with cancer? A survey of Australian oncologists. Support Care Cancer 2018; 26:451–60.

145. van Abbema DL, van den Akker M, Janssen-Heijnen ML, van den Berkmortel F, Hoeben A, de Vos-Geelen J et al. Patient- and tumor-related predictors of chemotherapy intolerance in older patients with cancer: a systematic review. J Geriatr Oncol 2018; 10:31–41.

146. Peterson LL, Hurria A, Feng T, Mohile SG, Owusu C, Klepin HD et al. Association between renal function and chemotherapy-related toxicity in older adults with cancer. J Geriatr Oncol 2017; 8:96–101.

147. Maggiore RJ, Dale W, Gross CP, Feng T, Tew WP, Mohile SG et al. Polypharmacy and potentially inappropriate medication use among older adults with cancer undergoing chemotherapy: impact on chemotherapy-related toxicity and hospitalization during treatment. J Am Geriatr Soc 2014; 62:1505–12.

148. Nightingale G, Hajjar E, Guo K, Komura S, Urnoski E, Sendecki J et al. A pharmacist-led medication assessment used to determine a more precise estimation of the prevalence of complementary and alternative medication (CAM) use among ambulatory senior adults with cancer. J Geriatr Oncol 2015; 6:411–7.

149. Alkan A, Yasar A, Karci E, Köksoy EB, Muslih Ü, Senler F et al. Severe drug interactions and potentially inappropriate medication usage in elderly cancer patients. Support Care Cancer 2017; 25:229–36.

150. Turner JP, Shakib S, Bell JS. Is my older cancer patient on too many medications? J Geriatr Oncol 2017; 8:77–81.

151. Guaraldo L, Cano FG, Damasceno GS, Rozenfeld S. Inappropriate medication use among the elderly: a systematic review of administrative databases. BMC Geriatr 2011; 11:1–10.

152. Saarelainen LK, Turner JP, Shakib S, Singhal N, Hogan-Doran J, Prowse R et al. Potentially inappropriate medication use in older people with cancer: Prevalence and correlates. J Geriatr Oncol 2014; 5:439–46.

153. Feng X, Higa GM, Safarudin F, Sambamoorthi U, Tan X. Potentially inappropriate medication use and associated healthcare utilization and costs among older adults with colorectal, breast, and prostate cancers. J Geriatr Oncol 2019; 10:698–704.

154. Zhang X, Zhou S, Pan K, Li X, Zhao X, Zhou Y et al. Potentially inappropriate medications in hospitalized older patients: a cross-sectional study using the Beers 2015 criteria versus the 2012 criteria. Clin Interv Aging 2017; 12:1697–703.

155. Sönnerstam E, Sjölander M, Gustafsson M. An evaluation of the prevalence of potentially inappropriate medications in older people with cognitive impairment living in Northern Sweden using the EU(7)-PIM list. Eur J Clin Pharmacol 2017; 73:735–42.

156. Motter FR, Fritzen JS, Hilmer SN, Paniz ÉV, Paniz VMV. Potentially inappropriate medication in the elderly: a systematic review of validated explicit criteria. Eur J Clin Pharmacol 2018; 74:679–700.

157. Morin L, Fastbom J, Laroche M-L, Johnell K. Potentially inappropriate drug use in older people: A nationwide comparison of different explicit criteria for population-based estimates. Br J Clin Pharmacol 2015; 80:315–24.

158. Curtin D, Gallagher PF, O'Mahony D. Explicit criteria as clinical tools to minimize inappropriate medication use and its consequences. Ther Adv Drug Saf 2019; 10:1–10.

159. Whitman A, DeGregory K, Morris A, Mohile S, Ramsdale E. Pharmacist-led medication assessment and deprescribing intervention for older adults with cancer and polypharmacy: a pilot study. Support Care Cancer 2018; 26:4105–13.

160. Karuturi MS, Holmes HM, Lei X, Johnson M, Barcenas CH, Cantor SB et al. Potentially inappropriate medications defined by STOPP criteria in older patients with breast and

colorectal cancer. J Geriatr Oncol 2019; 10:705–8.

161. Karuturi MS, Holmes HM, Lei X, Johnson M, Barcenas CH, Cantor SB et al. Potentially inappropriate medication use in older patients with breast and colorectal cancer. Cancer 2018; 124:3000–7.

162. Fulda TR, Valuck RJ, Zanden J V., Parker S, Byrns PJ. Disagreement among drug compendia on inclusion and ratings of drug-drug interactions. Curr Ther Res 2000; 61:540–8.

163. Nightingale G, Pizzi LT, Barlow A, Barlow B, Jacisin T, McGuire M et al. The prevalence of major drug-drug interactions in older adults with cancer and the role of clinical decision support software. J Geriatr Oncol 2018; 9:526–33.

164. Lund JL, Sanoff HK, Peacock-Hinton S, Muss H, Pate V, Stürmer T. Potential medication-related problems in older breast, colon, and lung cancer patients in the United States. Cancer Epidemiol Biomarkers Prev 2018; 27:41–9.

165. Leger DY, Moreau S, Signol N, Fargeas JB, Picat MA, Penot A et al. Polypharmacy, potentially inappropriate medications and drug-drug interactions in geriatric patients with hematologic malignancy: Observational single-center study of 122 patients. J Geriatr Oncol 2018; 9:60–7.

166. Zhou A, Holmes HM, Hurria A, Wildes TM. An analysis of the inclusion of medications considered potentially inappropriate in older adults in chemotherapy templates for hematologic malignancies: One recommendation for all? Drugs Aging 2018; 35:459–65.

167. Park JW, Roh J-L, Lee S, Kim S-B, Choi S-H, Nam SY et al. Effect of polypharmacy and potentially inappropriate medications on treatment and posttreatment courses in elderly patients with head and neck cancer. J Cancer Res Clin Oncol 2016; 142:1031–40.

168. Sokol K, Knudsen J, Li M. Polypharmacy in older oncology patients and the need for an interdisciplinary approach to side-effect management. J Clin Pharm Ther 2007; 32:169–75.

169. Deliens C, Deliens G, Filleul O, Pepersack T, Awada A, Piccart M et al. Drugs prescribed

for patients hospitalized in a geriatric oncology unit: Potentially inappropriate medications and impact of a clinical pharmacist. J Geriatr Oncol 2016; 7:463–70.

170. Colombo LRP, Aguiar PM, Lima TM, Storpirtis S. The effects of pharmacist interventions on adult outpatients with cancer: A systematic review. J Clin Pharm Ther 2017; 42:414–24.

8 Appendix

Appendix A – Study materials

- Informed consent form
- Patient information brochure

Appendix B – Risk assessment

- CARG score
- CRASH score
- IADL
- MMSE
- MNA
- CRASH scoring for chemotherapy toxicity (MAX2 Index) by analogy

Appendix C – Outcome measurement

- CTCAE documentation form
- PRO-CTCAE documentation form

Appendix D – Results of the evaluation study

- Distribution of CTCAE grade ≥ 3 toxicity during therapy course per individual CARG and CRASH score items and patient characteristics
- Incidence of severe toxicity at start of therapy (first cycle or at least 3 weeks of therapy)
- Relationship between the hematologic/nonhematologic CRASH score with the CARG score
- Relationship between the hematologic/nonhematologic CRASH score with the hematologic/nonhematologic physicians' judgments
- Comparison of the proportion of patients with hematologic toxicity per hematologic CRASH score value in our study with the original CRASH score development study
- Comparison of the proportion of patients with nonhematologic toxicity per nonhematologic CRASH score value in our study with the original CRASH score development study
- Kaplan-Meier estimates of the CARG score and combined CRASH score for severe overall toxicity
- Predictive performance of the ECOG performance status and age for severe hematologic and nonhematologic toxicity

- Risk factors of nonhematologic and hematologic CTCAE grade ≥ 3 toxicity during therapy course
- Predictive performance for the alterations of the planned treatment
- Kaplan-Meier analysis and Cox Regression for total and minor alterations of the planned treatment
- Kaplan-Meier estimates of the CARG score and the combined CRASH score for major alterations of the planned treatment (discontinuations or changes)

Appendix E – Results of the medication risk analysis

- Prevalence of antineoplastic agents and supportive care medication after start of cancer therapy
- Distribution of CTCAE grade ≥ 3 toxicity during therapy course regarding medication risks

Appendix A

Study materials

Informed consent form

DIE JOHANNITER. Aus Liebe zum Leben

universitätbonn

Version vom 23. September 2015

Rheinische Friedrich-Wilhelms-Universität Bonn

Pharmazeutisches Institut
Klinische Pharmazie

Einwilligungserklärung

Prof. Dr. Ulrich Jaehde
Imke Ortland

Ansprechpartnerin:
Imke Ortland
An der Immenburg 4
53121 Bonn
Tel.: 0228/735229
i.ortland@uni-bonn.de

Vorname, Name: _____

Geburtsdatum: _____

Ich erkläre, dass ich die Patienteninformation zur wissenschaftlichen Untersuchung

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

und diese Einwilligungserklärung in Kopie erhalten habe.

☐ Ich wurde ausreichend mündlich und schriftlich über die wissenschaftliche Untersuchung informiert.

☐ Ich weiß, dass ich jederzeit meine Einwilligung, ohne Angabe von Gründen, widerrufen kann, ohne dass dies für mich nachteilige Folgen hat. Es besteht die Möglichkeit, meine bereits erhobenen personenbezogenen Daten auf Wunsch zu löschen.

☐ **Ich bin damit einverstanden, dass die im Rahmen der wissenschaftlichen Untersuchung über mich erhobenen Krankheitsdaten, die durch Routineuntersuchungen erhaltenen Labordaten sowie meine sonstigen mit dieser Untersuchung zusammenhängenden personenbezogenen Daten von einer nicht ärztlichen Mitarbeiterin des Projektes aufgezeichnet werden. Dies gilt für die Daten des Johanniter Krankenhauses Bonn, sowie die der gegebenenfalls weiterbehandelnden onkologischen Praxis. Zu diesem Zweck entbinde ich, soweit dies für die Projektdurchführung notwendig ist, meine behandelnden Ärzte des Johanniter Krankenhauses Bonn und der gegebenenfalls weiterbehandelnden onkologischen Praxis von ihrer ärztlichen Schweigepflicht. Es wird gewährleistet, dass meine personenbezogenen Daten und meine Patientenakte nicht an Dritte weitergegeben werden. Zu diesem Zwecke willige ich ein, dass diese Daten im Rahmen der wissenschaftlichen Untersuchung verschlüsselt und gespeichert werden. Bei der Veröffentlichung in einer wissenschaftlichen Zeitschrift wird aus den Daten nicht hervorgehen, wer an dieser Untersuchung teilgenommen hat. Meine persönlichen Daten unterliegen dem Datenschutzgesetz.**

☐ Mit der vorstehend geschilderten Vorgehensweise bin ich einverstanden und bestätige dies mit meiner Unterschrift.

_____, den _____ _____
Ort Datum Unterschrift

Imke Ortland

_____ _____
Name wissenschaftliche Mitarbeiterin Unterschrift

Patient information brochure

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

-Beobachtungsstudie-

Patienteninformation

Version vom 23. September 2015

Verantwortliche Leiter:
Prof. Dr. A. H. Jacobs, Geriatrie mit Neurologie und Tagesklinik,
Johanniter Krankenhaus Bonn
Prof. Dr. Y.-D. Ko, Innere Medizin, Hämatologie und Onkologie,
Johanniter Krankenhaus Bonn
Prof. Dr. U. Jaehde, Klinische Pharmazie, Universität Bonn

Kontaktadresse:
Imke Ortland
Pharmazeutisches Institut
Klinische Pharmazie
An der Immenburg 4
53121 Bonn

E-Mail: i.ortland@uni-bonn.de

Telefon: 0228/73-5229

Sehr geehrte Patientin, sehr geehrter Patient,

im Rahmen einer Kooperation zwischen dem Johanniter Krankenhaus Bonn und dem Bereich Klinische Pharmazie der Universität Bonn möchten wir in einer Studie untersuchen, ob ältere Patienten von einem so genannten „onkogeriatrischen Assessment“ profitieren. Hierbei wird anhand von festgelegten Kriterien eine Abschätzung des Risikos einer gegebenenfalls notwendigen Tumortherapie vorgenommen. Ein „onkogeriatrisches Assessment“ steht hierbei für die Beurteilung (englisch „assessment“) der persönlichen Situation eines Patienten vor dem Hintergrund seines Lebensalters (Geriatrie) und der Tumorbehandlung (Onkologie).

Dabei sind wir auf Ihre Hilfe angewiesen.

In dem Ihnen vorliegenden Informationsmaterial wird Ihnen das geplante Projekt genau vorgestellt. Es wird beschrieben, welche Überlegungen zur Planung des Projektes geführt haben, wie das Projekt ablaufen soll und was eine Teilnahme für Sie als Patient ganz praktisch bedeuten würde.

Nehmen Sie sich für das Lesen ruhig viel Zeit. Legen Sie die Unterlagen zwischendurch beiseite, um darüber nachzudenken. Machen Sie sich überall in dieser Information Notizen zu den Dingen, die Sie gerne noch mit uns klären würden.

Sollte Ihnen während des Lesens irgendetwas unklar erscheinen, so scheuen Sie sich nicht, Ihren behandelnden Arzt oder die verantwortliche wissenschaftliche Mitarbeiterin anzusprechen.

Wir danken Ihnen im Voraus für Ihr Interesse an dieser Studie.

Prof. Dr. Andreas Jacobs
(Projektleiter)

Prof. Dr. Ulrich Jaehde
(Projektleiter)

Prof. Dr. Yon-Dschun Ko
(Projektleiter)

Imke Ortland
(wissenschaftliche Mitarbeiterin)

Warum ist ein „onkogeriatrisches Assessment“ im Alter sinnvoll?

Je älter ein Mensch wird, desto höher ist das Risiko an Krebs zu erkranken. Viele ältere Patienten können heutzutage eine Chemotherapie erhalten und davon profitieren. Eine Chemotherapie richtet sich gegen Zellen in Ihrem Körper, die sich häufig teilen und sich dadurch erneuern. Die verwendeten Arzneistoffe können nicht zwischen kranken und gesunden Zellen unterscheiden. Dies kann zu Nebenwirkungen führen. Zusätzlich kann der Körper im Alter anders auf Arzneistoffe reagieren. Das bedeutet, dass eine erhöhte Sensibilität gegenüber bestimmten Arzneistoffen vorliegen kann. Diese Gegebenheiten erschweren die Einschätzung des Arztes, ob ein älterer Patient von einer Chemotherapie profitieren wird oder ob eventuell auftretende Nebenwirkungen überwiegen.

Um dem Arzt die Einschätzung des Therapierisikos einer Chemotherapie bei älteren Patienten zu erleichtern, möchten wir in diesem Projekt in Form eines „onkogeriatrischen Assessments“ zwei Punkte-Systeme testen, welche die Verträglichkeit einer Chemotherapie bei älteren Patienten strukturiert abschätzen. Ziel ist es, den Nutzen und das Risiko einer Chemotherapie in Zukunft besser beurteilen zu können. Da bisher noch nicht ausreichend bekannt ist, ob ein „onkogeriatrisches Assessment“ die Sicherheit des Patienten während der Therapie tatsächlich erhöhen kann, wurde diese Studie initiiert.

Was bedeutet eine Teilnahme an dieser Studie konkret für Sie?

Eine Teilnahme an dieser Studie ist für Sie selbstverständlich freiwillig. Wir möchten Sie dennoch herzlich einladen, an diesem Projekt teilzunehmen.

Wenn Sie einer Teilnahme zustimmen, wird bei Ihnen ein „onkogeriatrisches Assessment“ durchgeführt. Hierzu werden mittels Ihrer Daten die beiden Punkte-Systeme erhoben. Die Ergebnisse der Punkte-Systeme werden die ärztliche Therapieauswahl nicht beeinflussen, denn es soll untersucht werden, ob die Einschätzung des Arztes oder die Aussage der Punkte-Systeme besser mit den möglicherweise tatsächlich auftretenden Nebenwirkungen übereinstimmt. Das „onkogeriatrische Assessment“ findet unter Einsatz der beiden Punkte-Systeme nur ein einziges Mal zu Beginn der Studie statt.

Um die Aussagekraft des onkogeriatrischen Assessments beurteilen zu können, werden bei Ihnen anschließend im Therapieverlauf gegebenenfalls auftretende Nebenwirkungen erfasst. Sie werden dazu einmal im Therapiezyklus gebeten, einen kurzen Fragebogen zu beantworten. Weiterhin wird die wissenschaftliche Mitarbeiterin ihre Patientenakte auf das Auftreten von weiteren Nebenwirkungen hin überprüfen.

Die Teilnahme an der Studie endet für Sie nach dem letzten oder spätestens nach dem sechsten Zyklus Ihrer verordneten Therapie. Sollten Sie keine medikamentöse Therapie erhalten, endet für Sie die Teilnahme nach der Erhebung der beiden Punkte-Systeme. Natürlich können Sie die Teilnahme auch jederzeit abbrechen, wenn Sie dies wünschen. In diesem Fall entstehen Ihnen keinerlei Nachteile hinsichtlich Ihrer Therapie. Es besteht die Möglichkeit, Ihre bereits erhobenen personenbezogenen Daten auf Wunsch zu löschen.

Welche Daten werden von Ihnen erhoben?

Die Information, die Sie bisher über diese Studie erhalten haben, lässt schon vermuten, dass eine Vielzahl von Daten über Ihre Person erfasst werden sollen. Dies geschieht allerdings erst dann, wenn dazu Ihre schriftliche Einwilligung vorliegt.

Persönliche Daten

Im Rahmen dieser Studie werden Ihre persönlichen Daten (z.B. Alter, Geschlecht, Gewicht, Größe), Informationen zu Therapie und Krankheit (z.B. Diagnosen, Therapiewechsel oder -modifikationen), sowie vorhandene Begleiterkrankungen erfasst. Ebenso werden für die Einschätzung Ihrer Erkrankung und auftretenden Nebenwirkungen relevante Daten aus der vom Arzt des Johanniter Krankenhauses Bonn oder vom Arzt der weiterbehandelnden onkologischen Praxis geführten

Patientenakte entnommen (z.B. Laborwerte). Auch Vitaminpräparate, Spurenelemente (z.B. Zink, Selen), Mineralstoffe (z.B. Calcium, Magnesium) sowie Arzneimittel zur Therapie anderer Krankheiten werden hierbei miterfasst.

Onkogeriatrisches Assessment

Innerhalb der beiden Punkte-Systeme werden verschiedene Parameter, beispielsweise Alter, Blutwerte, Gewicht und ausführliche Fragen zu Ihrer Gesundheit oder zu Ihren täglichen und sozialen Aktivitäten zusammen als Punkte-System analysiert. Die Laborwerte werden dabei Ihrer Krankenakte sowie vorhandenen Laborausdrucken entnommen. Mit dem resultierenden Wert dieser Punktzahl lässt sich die Verträglichkeit einer Chemotherapie im Voraus abschätzen.

Das hierzu erforderliche Gespräch findet während Ihres üblichen Aufenthaltes im Johanniter Krankenhaus statt und hat eine voraussichtliche Zeitdauer von 20-30 Minuten.

Fragebogen zur Erfassung von Nebenwirkungen und Lebensqualität

Um die bei Ihnen tatsächlich aufgetretenen Nebenwirkungen zu erfassen, werden Sie gebeten, den Schweregrad bestimmter Nebenwirkungen mittels eines Fragebogens einzustufen. Hierzu befragen wir Sie einmal pro Zyklus während Ihres üblichen Aufenthaltes im Johanniter Krankenhaus oder telefonisch. Außerdem bitten wir Sie im 4. Zyklus einmalig anhand einiger Zusatzfragen Ihre aktuelle Lebensqualität einzuschätzen.

Diese Fragebögen können innerhalb von fünf bis zehn Minuten entweder während Ihres üblichen Aufenthaltes im Johanniter Krankenhaus oder gegebenenfalls telefonisch beantwortet werden.

Information zum Datenschutz

Die gewonnenen Daten unterliegen den Bestimmungen des Datenschutzes und werden entsprechend der geltenden gesetzlichen Vorschriften pseudonymisiert. Das bedeutet, dass jeder teilnehmende Patient eine Patientennummer erhält und somit Ihr Name nicht in der Dokumentation und auf den Fragebögen erscheint. Es werden nur solche persönlichen Daten erhoben, mit denen ein Rückschluss auf Ihre Person oder Ihre Krankheit nicht oder nur mit unverhältnismäßig großem Aufwand an Zeit, Kosten und Arbeitskraft möglich ist. Die Daten werden von einer nicht ärztlichen Mitarbeiterin des Projektes ausschließlich zum Zweck der Durchführung der Studie erhoben und ausgewertet.

Daher bitten wir Sie, die Sie behandelnden Ärzte des Johanniter Krankenhauses, sowie gegebenenfalls der weiterbehandelnden onkologischen Praxis, von der ärztlichen Schweigepflicht zu entbinden, soweit dies für die Projektdurchführung notwendig ist. Bevor die Dokumentation beginnen kann, müssen Sie der Verwendung Ihrer Daten für Studienzwecke zustimmen. Es ist gewährleistet, dass aus Veröffentlichungen der in dieser Studie erhobenen Daten Ihr Name nicht hervorgeht. Die Ergebnisse der Studie werden anonymisiert veröffentlicht und stehen Ihnen dann selbstverständlich auf Anfrage zur Verfügung.

Die Teilnahme an dieser Studie beinhaltet für Sie keine zusätzlichen Risiken. In jedem Fall helfen Sie durch Ihre Teilnahme an dieser Studie, die Arzneimitteltherapie für Krebspatienten in Zukunft sicherer zu machen.

Eine Teilnahme an dieser Studie ist selbstverständlich freiwillig. Sie haben jederzeit das Recht ohne Angabe von Gründen von der Teilnahme zurückzutreten. In diesem Fall besteht die Möglichkeit, Ihre bereits erhobenen personenbezogenen Daten auf Wunsch zu löschen. Es entstehen Ihnen dadurch keine Nachteile hinsichtlich Ihrer Therapie!

Wenn Sie dieses Informationsmaterial eingehend gelesen haben und Ihre Fragen beantwortet wurden, können Sie frei über die Teilnahme an der Studie entscheiden. Ihre Teilnahme und Ihr Einverständnis mit den erläuterten Bestimmungen zum Datenschutz bestätigen Sie bitte schriftlich mit der **Einwilligungserklärung**, die Sie anbei erhalten haben.

Für Ihre Mühe und Ihre Mitarbeit danken wir Ihnen sehr herzlich!

Kontaktadresse

Imke Ortland

Pharmazeutisches Institut
Klinische Pharmazie
An der Immenburg 4
53121 Bonn

Appendix B

Risk assessment

CARG score

universitätbonn

CARG – Chemotherapie Toxizitätskalkulator

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum	
Parameter		
1. Geschlecht (Akte)	☐ männlich ☐ weiblich	
2. Alter (Akte)	_____ Jahre	
3. Größe (Akte)	_____ cm	
4. Gewicht (Akte)	_____ kg	
5. Krebsart (Akte)	☐ Gastrointestinal ☐ Urogenital ☐ Andere	Krebsart:
6. Dosierung (Akte)	☐ Standard ☐ Reduziert	
7. Anzahl Chemotherapeutika (Akte)	☐ Polychemotherapie ☐ Monochemotherapie	Anzahl an Chemotherapeutika:
8. Hämoglobin (Akte)	☐ < 10 g/dL (w) / <11g/dL (m) ☐ ≥ 10 g/dL (w) / ≥ 11g/dL (m)	Wert:
9. Wie ist Ihr Hörvermögen (mit einer Hörhilfe, wenn nötig)?	☐ Sehr gut ☐ Gut ☐ Ausreichend ☐ Schlecht ☐ Vollkommen taub	
10. Wie oft sind Sie in den letzten 6 Monaten gestürzt?	☐ ≥ 1 ☐ 0	Anzahl Stürze:
11. Können Sie Ihre eigenen Medikamente einnehmen?	☐ Ohne Hilfe (richtige Dosierung zur richtigen Tageszeit); oder ☐ Mit etwas Hilfe (Sie können Ihre Medikamente einnehmen, wenn jemand sie für Sie vorbereitet und/oder Sie daran erinnert); oder ☐ Sie können Ihre Medikamente nicht mehr eigenständig einnehmen	
12. Schränkt Ihr Gesundheitszustand Sie ein, wenn Sie eine kurze Strecke gehen?	☐ Sehr eingeschränkt ☐ Ein bisschen eingeschränkt ☐ Überhaupt nicht eingeschränkt	
13. Während der letzten 4 Wochen: wie oft haben Ihr körperlicher Gesundheitszustand oder Ihre seelischen Probleme Ihre sozialen Aktivitäten beeinträchtigt (z.B. sich mit Freunden oder Verwandten treffen, usw.)?	☐ Immer ☐ Meistens ☐ Manchmal ☐ Selten ☐ Nie	
14. Serumkreatinin (Akte):	_____ mg/dL	
Ergebnisse		
Kreatininclearance (Jelliffe)		
Toxizitätsscore		
Risiko Chemotherapie Toxizität	_____ %	
Toxizitätskategorie		

CRASH score

universitätbonn

CRASH – Score

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum

Risiko der Chemotherapie (nicht vorhandenen Regime - Analogie) – Regime unterstreichen:			Scorepunkzahl
☐ 0 Capecitabin 2 g; Cisplatin/Pemetrexed; Dacarbazin; Docetaxel wöchentlich; FOLFIRI; Gemcitabin 1 g; 3/4 Wochen; Gemcitabin 1,25 g 3/4 Wochen; Paclitaxel wöchentlich; Pemetrexed	☐ 1 Bendamustin +/- Rituximab; Capecitabin 2,5 g; Carboplatin/Gemcitabin; AUC 4-6/1 g d1, d8; Carboplatin/Pemetrexed; Cisplatin/Gemcitabin d1,8; ECF; Fludarabin; FOLFOX 85 mg (z.B. FOLFOX4 oder mFOLFOX6); Gemcitabin 7/8 Wochen, dann 3/4; Gemcitabin/Irinotecan; PEG Doxorubicin 50 mg q4w; Topotecan wöchentlich; XELOX	☐ 2 5-FU/LV (Roswell-Park); 5-FU/LV (Mayo); 5-FU/LV + Bevacizumab; AC; CAF; Carboplatin/Doce oder Paclitaxel q3w; CHOP +/- Rituximab; Cisplatin/Docetaxel 75/75; Cisplatin/Etoposid; Cisplatin/Gemcitabin d1, d8, d15; Cisplatin/Irinotecan q3w; Cisplatin/Paclitaxel 135-24 h q3w; CMF classic,; Doxorubicin q3w; FOLFOX 100-130 mg; Gemcitabin/Pemetrexed d8;Irinotecan q3w; Paclitaxel q3w; Docetaxel q3w; Topotecan monatlich	

Hämatologische Risikofaktoren			
Diastolischer Blutdruck	Wert:	☐ > 72 = 1 ☐ sonst = 0	
IADL	Punktzahl:	☐ < 26 = 1 ☐ sonst = 0	
LDH (Laktatdehydrogenase)	Wert:	☐ > 167 =2 (> 0,75 ULN) ☐ sonst = 0	
Nicht-Hämatologische Risikofaktoren			
ECOG – PS	Punktzahl:	☐ 0 = 0 ☐ 1-2 = 1 ☐ 3-4 = 2	
MMS	Punktzahl:	☐ < 30 = 2 ☐ 30 = 0	
MNA	Punktzahl:	☐ < 28 = 2 ☐ sonst = 0	

CRASH – Risiko Score

Hämatologisch	Nicht-Hämatologisch	Kombiniert	Kategorie
0 – 1: 7 %	0 – 2: 33 %	0 – 3: 50 %	Low
2 – 3: 23 %	3 – 4: 46 %	4 – 6: 58 %	Int-Low
4 – 5: 54 %	5 – 6: 67 %	7 – 9: 77 %	Int-High
≥ 5: 100 %	≥ 6: 93 %	≥ 9: 79 %	High

		Score	Risiko in %	Risiko-Kategorie
CRASH	Hämatologisches Risiko			
	Nicht-Hämatologisches Risiko			
	Kombiniertes Risiko			

IADL (Instrumental activities of daily living)

universitätbonn

CRASH – IADL

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum
	Ergebnis /28

IADL (Lawton 1988)

1. Können Sie das Telefon benutzen:
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können das Telefon überhaupt nicht mehr benutzen?

2. Können Sie Orte besuchen, die für Sie fußläufig nicht mehr erreichbar sind?
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können überhaupt nicht mehr reisen, außer unter speziellen Vorkehrungen?

3. Können Sie Lebensmittel einkaufen:
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können überhaupt nicht mehr einkaufen gehen?

4. Können Sie Ihre eigenen Mahlzeiten zubereiten:
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können überhaupt nicht mehr eigene Mahlzeiten zubereiten?

5. Können Sie den Haushalt machen:
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können überhaupt nicht mehr den Haushalt machen?

6. Können Sie handwerkliche Arbeiten erledigen:
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können überhaupt nicht mehr handwerkliche Arbeiten erledigen?

7. Können Sie Ihre Wäsche machen:
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können die Wäsche überhaupt nicht mehr machen?

8. Nehmen Sie Medikamente ein:
   - 1 ☐ Ja (8b)
   - 2 ☐ Nein (8c)

8b. Können Sie Ihre eigene Medikamente einnehmen:
   - 3 ☐ Ohne Hilfe (In der richtigen Dosierung und zur richtigen Tageszeit)
   - 2 ☐ Mit etwas Hilfe (Sie können die Medikamente einnehmen, wenn jemand diese für Sie vorbereitet und/oder Sie an die Einnahme erinnert)
   - 1 ☐ Sie können überhaupt nicht mehr Ihre eigenen Medikamente einnehmen?

8c. Wenn Sie früher Medikamente eingenommen haben, konnten Sie diese einnehmen:
   - 3 ☐ Ohne Hilfe (In der richtigen Dosierung und zur richtigen Tageszeit)
   - 2 ☐ Mit etwas Hilfe (Sie können die Medikamente einnehmen, wenn jemand diese für Sie vorbereitet und/oder Sie an die Einnahme erinnert)
   - 1 ☐ Sie könnten überhaupt nicht mehr Ihre eigenen Medikamente einnehmen?

9. Können Sie Ihre Geldgeschäfte selbständig erledigen:
   - 3 ☐ Ohne Hilfe
   - 2 ☐ Mit etwas Hilfe
   - 1 ☐ Sie können überhaupt nicht mehr Ihre Geldgeschäfte erledigen?

MMSE (Mini-Mental State Examination)

universitätbonn

CRASH – MMS

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum
	Ergebnis /30

Mini-Mental State Examination (MMSE)

(0 / 1)	1.	Was für ein Datum ist heute?	
(0 / 1)	2.	Welche Jahreszeit?	
(0 / 1)	3.	Welches Jahr haben wir?	
(0 / 1)	4.	Welcher Wochentag ist heute?	
(0 / 1)	5.	Welcher Monat?	
(0 / 1)	6.	Wo sind wir jetzt ?	welches Bundesland?
(0 / 1)	7.		welcher Landkreis/welche Stadt?
(0 / 1)	8.		welche Stadt/welcher Stadtteil?
(0 / 1)	9.		welches Krankenhaus?
(0 / 1)	10.		welche Station/welches Stockwerk?
(0 / 1)	11.	Bitte merken Sie sich:	Apfel
(0 / 1)	12.		Pfennig
(0 / 1)	13.		Tisch
			Anzahl der Versuche:

Ziehen Sie von 100 jeweils 7 ab oder buchstabieren Sie Stuhl rückwärts:

(0 / 1)	14.	93	L
(0 / 1)	15.	86	H
(0 / 1)	16.	79	U
(0 / 1)	17.	72	T
(0 / 1)	18.	65	S

Was waren die Dinge, die Sie sich vorher gemerkt haben ?

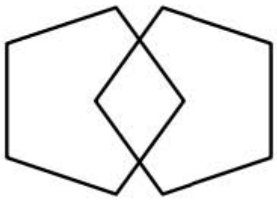
(0 / 1)	19.		Apfel
(0 / 1)	20.		Pfennig
(0 / 1)	21.		Tisch
(0 / 1)	22.	Was ist das?	Uhr
(0 / 1)	23.		Bleistift/Kugelschreiber
(0 / 1)	24.	Sprechen Sie nach:	"Kein wenn und oder aber."

Machen Sie bitte folgendes:

(0 / 1)	25.	Nehmen Sie bitte das Blatt in die Hand,
(0 / 1)	26.	Falten Sie es in der Mitte und
(0 / 1)	27.	Lassen Sie es auf den Boden fallen
(0 / 1)	28.	Lesen Sie und machen Sie es bitte ("**Augen zu** !")
(0 / 1)	29.	Schreiben Sie bitte einen Satz (mind. Subjekt und Prädikat)
(0 / 1)	30.	Kopieren Sie bitte die Zeichnung (zwei Fünfecke)

[1] Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie. Mini-Mental State Examination. Verfügbar unter: www.dgho.de. Letzter Zugriff 7. Januar 2015.

Augen zu!



MNA (Mini Nutritional Assessment)

universitätbonn

CRASH – MNA
Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum
	Ergebnis /30

Mini Nutritional Assessment MNA®

Nestlé NutritionInstitute

Name: Vorname:

Geschlecht: Alter (Jahre): Gewicht (kg): Größe (m): Datum:

Füllen Sie den Bogen aus, indem Sie die zutreffenden Zahlen in die Kästchen eintragen. Addieren Sie die Zahlen des Screenings. Ist der Wert ≤ 11, fahren Sie mit dem Assessment fort, um den Mangelernährungs-Index zu erhalten.

Screening

A Hat der Patient während der letzten 3 Monate wegen Appetitverlust, Verdauungsproblemen, Schwierigkeiten beim Kauen oder Schlucken weniger gegessen?
0 = starke Abnahme der Nahrungsaufnahme
1 = leichte Abnahme der Nahrungsaufnahme
2 = keine Abnahme der Nahrungsaufnahme ☐

B Gewichtsverlust in den letzen 3 Monaten
0 = Gewichtsverlust > 3 kg
1 = nicht bekannt
2 = Gewichtsverlust zwischen 1 und 3 kg
3 = kein Gewichtsverlust ☐

C Mobilität
0 = bettlägerig oder in einem Stuhl mobilisiert
1 = In der Lage, sich in der Wohnung zu bewegen
2 = verlässt die Wohnung ☐

D Akute Krankheit oder psychischer Stress während der letzten 3 Monate?
0 = ja 2 = nein ☐

E Neuropsychologische Probleme
0 = schwere Demenz oder Depression
1 = leichte Demenz
2 = keine psychologischen Probleme ☐

F Body Mass Index (BMI): Körpergewicht (kg) / Körpergröße² (m²)
0 = BMI < 19
1 = 19 ≤ BMI < 21
2 = 21 ≤ BMI < 23
3 = BMI ≥ 23 ☐

Ergebnis des Screenings (max. 14 Punkte) ☐☐

12-14 Punkte: Normaler Ernährungszustand
8-11 Punkte: Risiko für Mangelernährung
0-7 Punkte: Mangelernährung

Für ein tiefergehendes Assessment fahren Sie bitte mit den Fragen G-R fort

Assessment

G Lebt der Patient eigenständig zu Hause?
1 = ja 0 = nein ☐

H Nimmt der Patient mehr als 3 verschreibungspflichtige Medikamente pro Tag?
0 = ja 1 = nein ☐

I Hat der Patient Druck- oder Hautgeschwüre?
0 = ja 1 = nein ☐

J Wie viele Hauptmahlzeiten isst der Patient pro Tag?
0 = 1 Mahlzeit
1 = 2 Mahlzeiten
2 = 3 Mahlzeiten ☐

K Eiweißzufuhr: Isst der Patient
- mindestens einmal pro Tag Milchprodukte (Milch, Käse, Joghurt)? ja☐ nein☐
- mindestens zweimal pro Woche Hülsenfrüchte oder Eier? ja☐ nein☐
- täglich Fleisch, Fisch oder Geflügel? ja☐ nein☐

0,0 = wenn 0 oder 1 mal «ja»
0,5 = wenn 2 mal «ja»
1,0 = wenn 3 mal «ja» ☐,☐

L Isst der Patient mindestens zweimal pro Tag Obst oder Gemüse?
0 = nein 1 = ja ☐

M Wie viel trinkt der Patient pro Tag? (Wasser, Saft, Kaffee, Tee, Milch ...)
0,0 = weniger als 3 Gläser / Tassen
0,5 = 3 bis 5 Gläser / Tassen
1,0 = mehr als 5 Gläser / Tassen ☐,☐

N Essensaufnahme mit / ohne Hilfe
0 = braucht Hilfe beim Essen
1 = isst ohne Hilfe, aber mit Schwierigkeiten
2 = isst ohne Hilfe, keine Schwierigkeiten ☐

O Wie schätzt der Patient seinen Ernährungszustand ein?
0 = mangelernährt
1 = ist sich unsicher
2 = gut ernährt ☐

P Im Vergleich mit gleichaltrigen Personen schätzt der Patient seinen Gesundheitszustand folgendermaßen ein:
0,0 = schlechter
0,5 = weiß es nicht
1,0 = gleich gut
2,0 = besser ☐,☐

Q Oberarmumfang (OAU in cm)
0,0 = OAU < 21
0,5 = 21 ≤ OAU ≤ 22
1,0 = OAU > 22 ☐,☐

R Wadenumfang (WU in cm)
0 = WU < 31
1 = WU ≥ 31 ☐

Assessment (max. 16 Punkte) ☐☐,☐
Screening ☐☐,☐
Gesamtauswertung (max. 30 Punkte) ☐☐,☐

Auswertung des Mangelernährungs-Index

24-30 Punkte ☐ Normaler Ernährungszustand
17-23,5 Punkte ☐ Risiko für Mangelernährung
Weniger als 17 Punkte ☐ Mangelernährung

Ref. Vellas B, Villars H, Abellan G, et al. *Overview of MNA® - Its History and Challenges.* J Nut Health Aging 2006; 10: 456-485.
Rubenstein LZ, Harker JO, Salva A, Guigoz Y, Vellas B. Screening for Undernutrition in Geriatric Practice: *Developing the Short-Form Mini Nutritional Assessment (MNA-SF).* J. Geront 2001; 56A: M366-377.
Guigoz Y. The Mini-Nutritional Assessment (MNA®) *Review of the Literature – What does it tell us?* J Nutr Health Aging 2006; 10: 466-487.

Mehr Informationen unter: www.mna-elderly.com

[1] MNA. Mini Nutritional Assessment. Nestlé Nutrition Institute. Verfügbar unter: www.mna-elderly.com. Letzter Zugriff 7. Januar 2015.

CRASH scoring for chemotherapy toxicity (MAX2 Index) by analogy

Score points were derived by analogy as recommended [68] or by the supplementary MAX2 list provided by Extermann to the author of this thesis

Therapy regimen*	**CRASH score points**
Rituximab-Mini-CHOP (cyclophosphamide, doxorubicin, vincristine, prednisone)	1
Carboplatin AUC2 + etoposide	1
Carboplatin AUC2 + paclitaxel weekly	1
Docetaxel + cyclophosphamide	2
Decitabine	0
Rituximab	0
Afatinib	0
Trastuzumab	0
Cisplatin + vinorelbine	2
Carboplatin AUC2 + paclitaxel 50-100 mg/m^2	1
Gemcitabine + Nab-paclitaxel	2
Carboplatin AUC2 + gemcitabine 800 mg/m^2	0
Carboplatin AUC2 + gemcitabine 1000 mg/m^2	0
Carboplatin AUC2 weekly	0
Doxorubicin 60 mg/m^2 + cisplatin 50 mg/m^2 3-weekly	2
Fluorouracil low dose	0
FLO (fluorouracil, oxaliplatin, folinic acid)	1
FUFOX (fluorouracil, oxaliplatin, folinic acid)	1
Rituximab-IMVP (ifosfamide, methotrexate, etoposide)-16	1
Paclitaxel weekly + trastuzumab	0
mFOLFOX6 (fluorouracil, oxaliplatin, folinic acid) + panitumumab	1
VMP (bortezomib, melphalan, prednisone)	0
Bortezomib + dexamethasone	0
Rituximab + gemcitabine + oxaliplatin	2
Paclitaxel + gemcitabine + oxaliplatin	2
Paclitaxel 90 mg/m^2 + bevacizumab	1
Doxorubicin 25 mg/m^2 + cyclophosphamide 500 mg/m^2 3-weekly	1
Pembrolizumab	0

Therapy regimen*	**CRASH score points**
Rituximab + chlorambucil	0
Lenalidomide + dexamethasone	0
Docetaxel + cyclophosphamide	2
Carboplatin AUC6 + etoposide 3-weekly	2
VCD (bortezomib, cyclophosphamide, dexamethasone)	1
Cytarabine	0
Epirubicine + cyclophosphamide	2
Ruxolitinib	0
Azacitidine	0
Capecitabine 2 g + trastuzumab	0
Cisplatin 50 mg/m^2 + doxorubicin 60 mg/m^2	2
Rituximab + chlorambucil	0
Ipilimumab	2
Idarubicin + tretinoin	2
Bevacizumab + topotecan + paclitaxel	2

*if no dose is specified, all common doses used for this regimen fall into the same category [68]

Appendix C

Outcome measurement

CTCAE documentation form

universitätbonn

Erhebung Toxizität

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum	Zyklus (Zeitraum)

Art der Toxizität	Parameter	Grad 1	Grad 2	Grad 3	Grad 4
Hämatologisch	Anämie (Reduziertes Hämoglobin)	Female <11.2 g/dL- 10.0 g/dL Male <13.7g/dL – 10.0 g/dL	Hb <10.0 - 8.0 g/dL	Hb <8.0 g/dL; transfusion indicated	Life-threatening consequences; urgent intervention indicated
Hämatologisch	Febrile Neutropenie	-	-	ANC <1.0 x 10^9/L with a single temperature of >38.3 °C or a sustained temperature of ≥38°C for more than one hour	Life-threatening consequences; urgent intervention indicated
Hämatologisch	Reduzierte Neutrophilenzahl	Female <1.56 x 10^9/L – 1.5 x 10^9/L Male <1.78 x 10^9/L – 1.5 x 10^9/L	<1.5 – 1.0 x 10^9/L	<1.0 - 0.5 x 10^9/L	<0.5 x 10^9/L
Hämatologisch	Reduzierte Leukozytenzahl	Female <3.98 x 10^9/L - 3.0 x 10^9/L Male 4.23 x 10^9/L - 3.0 x 10^9/L	<3.0 - 2.0 x 10^9/L	<2.0 - 1.0 x 10^9 /L	<1.0 x 10^9 /L
Hämatologisch	Reduzierte Thrombozytenzahl	Female <182.0- 75.0 x 10^9 /L Male <163.0- 75.0 x 10^9 /L	<75.0 - 50.0 x 10^9 /L	<50.0 - 25.0 x 10^9 /L	<25.0 x 10^9/L
Kardiovaskulär	Akutes Koronarsyndrom	-	Symptomatic, progressive angina; cardiac enzymes normal; hemo-dynamically stable	Symptomatic, unstable angina and/or acute myocardial infarction, cardiac enzymes abnormal, hemo-dynamically stable	Symptomatic, unstable angina and/or acute myocardial infarction, cardiac enzymes abnormal, hemo-dynamically unstable
Kardiovaskulär	Herzflattern	Asymptomatic, intervention not indicated	Non-urgent medical intervention indicated	Symptomatic and incompletely controlled medically, or controlled with device (e.g., pacemaker), or ablation	Life-threatening consequences; urgent intervention indicated
Kardiovaskulär	Herzinsuffizienz	Asymptomatic with laboratory (e.g. BNP or cardiac imaging abnormalities	Symptoms with mild to moderate activity or exertion	Severe with symptoms at rest or with minimal activity or exertion; intervention indicated	Life-threatening consequences; urgent intervention indicated (e.g., continuous IV therapy or mechanical hemodynamic support)

universitätbonn

Erhebung Toxizität

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Art der Toxizität	Parameter	Grad 1	Grad 2	Grad 3	Grad 4
Kardiovaskulär	Hypertonie	Prehypertension (systolic BP 120 - 139 mm Hg or diastolic BP 80 - 89 mm Hg)	Stage 1 hypertension (systolic BP 140 - 159 mm Hg or diastolic BP 90 - 99 mm Hg); medical intervention indicated; recurrent or persistent (≥24 hrs); symptomatic increase by >20 mm Hg (diastolic) or to >140/90 mm Hg if previously WNL; monotherapy indicated	Stage 2 hypertension (systolic BP ≥160 mm Hg or diastolic BP ≥100 mm Hg); medical intervention indicated; more than one drug or more intensive therapy than previously used indicated	Life-threatening consequences (e.g., malignant hypertension, transient or permanent neurologic deficit, hypertensive crisis); urgent intervention indicated Pediatric: Same as adult
Thrombo-embolie	Thrombo-embolisches Ereignis	Venous thrombosis (e.g., superficial thrombosis)	Venous thrombosis (e.g., uncomplicated deep vein thrombosis), medical intervention indicated	Thrombosis (e.g., uncomplicated pulmonary embolism, nonembolic cardiac mural thrombus), medical intervention indicated	Life-threatening (e.g., pulmonary embolism, cerebrovascular event, arterial insufficiency); hemodynamic or neurologic instability; urgent intervention indicated
Leberfunktion	Erhöhte GOT	Female >35 – 105 U/L Male >50 -150 U/L	Female 105 U/L- 175 U/L Male >150 U/L– 250 U/L	Female 175 U/L- 700 U/L Male >250 U/L – 1000 U/L	Female >700 U/L Male >1000 U/L
Leberfunktion	Erhöhte GPT	Female >35 – 105 U/L Male >50 -150 U/L	Female 105 U/L- 175 U/L Male >150 U/L– 250 U/L	Female 175 U/L- 700 U/L Male >250 U/L – 1000 U/L	Female >700 U/L Male >1000 U/L
Leberfunktion	Erhöhte GGT	Female >42 U/L – 105 U/L Male > 71 U/L – 177.5 U/L	Female >105 U/L – 210 U/L Male > 177.5 U/L – 355 U/L	Female >210 U/L- 840 U/L Male > 355 U/L-1420 U/L	Female >840 U/L Male > 1420 U/L
Leberfunktion	Erhöhtes Gesamt-Bilirubin	>1.2 mg/dL – 1.8 mg/dL	>1.8 mg/dL – 3.6 mg/dL	>3.6 mg/dL-12 mg/dl	>12 mg/dl
Niere	Erhöhtes Kreatinin	Female >0.9 mg/dl -1.35 mg/dl Male >1.2 mg/dl – 1.8 mg/dl	Female 1.35 mg/dl – 2.7 mg/dl Male >1.8 mg/dl – 3.6 mg/dl	Female 2.7 mg/dl- 5.4 mg/dl Male >1.8 mg/dl – 3.6 mg/dl	Female >5.4 mg/dl Male >3.6 mg/dl
Proteinurie	Proteinurie	1+ proteinuria; urinary protein <1.0 g/24 hrs	Adults: 2+ proteinuria; urinary protein 1.0 – 3.4 g/24 hrs	Adults: urinary protein >=3.5 g/24 hrs	-

universitätbonn

Erhebung Toxizität

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Art der Toxizität	Parameter	Grad 1	Grad 2	Grad 3	Grad 4
Atem-beschwerden	Dyspnoe	Shortness of breath with moderate exertion	Shortness of breath with minimal exertion; limiting instrumental ADL	Shortness of breath at rest; limiting self care ADL	Life-threatening consequences; urgent intervention indicated
Hautausschläge	Erythrodermie	-	Erythema covering >90% BSA without associated symptoms; limiting instrumental ADL	Erythema covering >90% BSA with associated symptoms (e.g., pruritus or tenderness); limiting self care ADL	Erythema covering >90% BSA with associated fluid or electrolyte abnormalities; ICU care or burn unit indicated
Hautausschläge	Urticaria	Urticarial lesions covering <10% BSA; topical intervention indicated	Urticarial lesions covering 10 - 30% BSA; oral intervention indicated	Urticarial lesions covering >30% BSA; IV intervention indicated	-
Hautausschläge	Hand-Fuß Syndrom	Minimal skin changes or dermatitis (e.g., erythema, edema, or hyperkeratosis) without pain	Skin changes (e.g., peeling, blisters, bleeding, edema, or hyperkeratosis) with pain; limiting instrumental ADL	Severe skin changes (e.g., peeling, blisters, bleeding, edema, or hyperkeratosis) with pain; limiting self care ADL	-
Blutungen	Magen- Darm-blutungen	Mild; intervention not indicated	Moderate symptoms; medical intervention or minor cauterization indicated	Transfusion, radiologic, endoscopic, or elective operative intervention indicated	Life-threatening consequences; urgent intervention indicated
Infektionen	Infektionen (verschiedene)	Asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated	Moderate; minimal, local or noninvasive intervention indicated; limiting age appropriate instrumental ADL	Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of existing hospitalization indicated; disabling; limiting self care ADL	Life-threatening consequences; urgent intervention indicated
Metabolische Störungen	Dehydration	Increased oral fluids indicated; dry mucous membranes; diminished skin turgor	IV fluids indicated <24 hrs	IV fluids or hospitalization indicated	Life-threatening consequences; urgent intervention indicated
Metabolische Störungen	Hyponatriämie	<136 mmol/L - 130 mmol/L	-	<130 - 120 mmol/L	<120 mmol/L; life-threatening consequences
Metabolische Störungen	Hypokaliämie	<3.5 mmol/L - 3.0 mmol/L	<3,5 mmol/L - 3.0 mmol/L; symptomatic; intervention indicated	<3.0 mmol/L- 2.5 mmol/L; hospitalization indicated	<2.5 mmol/L; life-threatening consequences

universitätbonn

Erhebung Toxizität

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum	Zyklus (Zeitraum)

Art der Toxizität	Parameter	Grad 1	Grad 2	Grad 3	Grad 4
Nicht-hämatologisch	**„Schwierigkeiten beim Schlucken“** →Dysphagia	Symptomatic, able to eat regular diet	Symptomatic and altered eating/ swallowing	Severely altered eating/swallowing; tube feeding or TPN or hospitalization indicated	Life-threatening consequences; urgent intervention indicated
Nicht-hämatologisch	**„Mundtrockenheit“** →Dry mouth	Symptomatic (e.g., dry or thick saliva) without significant dietary alteration; unstimulated saliva flow >0.2 ml/min	Moderate symptoms; oral intake alterations (e.g., copious water, other lubricants, diet limited to purees and/or soft, moist foods); unstimulated saliva 0.1 to 0.2 ml/min	Inability to adequately aliment orally; tube feeding or TPN indicated; unstimulated saliva <0.1 ml/min	-
Nicht-hämatologisch	**“Wunde/offene Stellen in Mund/Hals”** →Mucositis oral	Asymptomatic or mild symptoms; intervention not indicated	Moderate pain; not interfering with oral intake; modified diet indicated	Severe pain; interfering with oral intake	Life-threatening consequences; urgent intervention indicated
Nicht-hämatologisch	**„Schmerzen“** →Pain	Mild pain	Moderate pain; limiting instrumental ADL	Severe pain; limiting self care ADL	-
Nicht-hämatologisch	**„Appetitmangel“** →Anorexia	Loss of appetite without alteration in eating habits	Oral intake altered without significant weight loss or malnutrition; oral nutritional supplements indicated	Associated with significant weight loss or malnutrition (e.g., inadequate oral caloric and/or fluid intake); tube feeding or TPN indicated	Life-threatening consequences; urgent intervention indicated
Nicht-hämatologisch	**„Verstopfung“** →Constipation	Occasional or intermittent symptoms; occasional use of stool softeners, laxatives, dietary modification, or enema	Persistent symptoms with regular use of laxatives or enemas; limiting instrumental ADL	Obstipation with manual evacuation indicated; limiting self care ADL	Life-threatening consequences; urgent intervention indicated
Nicht-hämatologisch	„**Durchfall“** →Diarrhea	Increase of <4 stools per day over baseline; mild increase in ostomy output compared to baseline	Increase of 4 - 6 stools per day over baseline; moderate increase in ostomy output compared to baseline	Increase of >=7 stools per day over baseline; incontinence; hospitalization indicated; severe increase in ostomy output compared to baseline;limiting self care ADL	Life-threatening consequences; urgent intervention indicated

universitätbonn

Erhebung Toxizität

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Art der Toxizität	Parameter	Grad 1	Grad 2	Grad 3	Grad 4
Nicht-hämatologisch	**„Übelkeit“** →Nausea	Loss of appetite without alteration in eating habits	Oral intake decreased without sig. weight loss, dehydration or malnutrition	Inadequate oral caloric or fluid intake; tube feeding, TPN, or hospitalization indicated	
Nicht-hämatologisch	**„Erbrechen“** →Vomiting	1 - 2 episodes (separated by 5 minutes) in 24 hrs	3 - 5 episodes (separated by 5 minutes) in 24 hrs	>=6 episodes (separated by 5 minutes) in 24 hrs; tube feeding, TPN or hospitalization indicated	Life-threatening consequences; urgent intervention indicated
Nicht-hämatologisch	**“Probleme beim Schlafen”** → Insomnia	Mild difficulty falling asleep, staying asleep or waking up early	Moderate difficulty falling asleep, staying asleep or waking up early	Severe difficulty in falling asleep, staying asleep or waking up early	-
Nicht-hämatologisch	**“Erschöpfung”** → Fatigue	Fatigue relieved by rest	Fatigue not relieved by rest; limiting instrumental ADL	Fatigue not relieved by rest, limiting self care ADL	
Nicht-hämatologisch	**„Kurzatmigkeit“** →Dyspnea	Shortness of breath with moderate exertion	Shortness of breath with minimal exertion; limiting instrumental ADL	Shortness of breath at rest; limiting self care ADL	Life-threatening consequences; urgent intervention indicated
Nicht-hämatologisch	**“Taubheit/ Kribbeln”** →Peripheral sensory neuropathy	Asymptomatic; loss of deep tendon reflexes or paresthesia	Moderate symptoms; limiting instrumental ADL	Severe symptoms; limiting self care ADL	Life-threatening consequences; urgent intervention indicated

PRO-CTCAE documentation form

universitätbonn

PRO-CTCAE
Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum	Zyklus

PRO-CTCAE: Fragebogen zu Symptomen bei Krebspatienten mit Chemotherapie

Bitte beantworten Sie die folgenden Fragen zu häufig vorkommenden Symptomen während einer Krebstherapie selbst, indem Sie die Antwort ankreuzen, die für Sie am besten zutrifft. Bitte beziehen Sie sich bei der Antwort immer auf **die schwerste Ausprägung** des jeweiligen Symptoms **in den letzten 7 Tagen** (gemeint ist nicht der Durchschnittswert).

1. Während der letzten 7 Tage: wie STARK waren Ihre **Schwierigkeiten beim Schlucken** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

2. Während der letzten 7 Tage: wie STARK war Ihre **Mundtrockenheit** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

3. Während der letzten 7 Tage: wie STARK hatten Sie **wunde oder offene Stellen in Mund oder Hals** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

4. Während der letzten 7 Tage: wie SEHR haben **wunde oder offene Stellen in Mund oder Hals** Sie in ihren täglichen Aktivitäten GESTÖRT?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

5. Während der letzten 7 Tage: wie HÄUFIG hatten Sie **Schmerzen**?

Nie	Selten	Gelegentlich	Häufig	Fast immer
❑	❑	❑	❑	❑

6. Während der letzten 7 Tage: wie STARK waren Ihre **Schmerzen** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

7. Während der letzten 7 Tage: wie SEHR haben **Schmerzen** Sie in Ihren täglichen Aktivitäten GESTÖRT?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

8. Während der letzten 7 Tage: wie STARK war Ihr **Appetitmangel** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

universitätbonn

PRO-CTCAE

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum	Zyklus

9. Während der letzten 7 Tage: wie SEHR hat Ihr **Appetitmangel** Sie in Ihren täglichen Aktivitäten GESTÖRT?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

10. Während der letzten 7 Tage: wie STARK war Ihre **Verstopfung** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

11. Während der letzten 7 Tage: wie HÄUFIG hatten Sie **Durchfall**?

Nie	Selten	Gelegentlich	Häufig	Fast immer
❑	❑	❑	❑	❑

12. Während der letzten 7 Tage: wie HÄUFIG hatten Sie **Übelkeit**?

Nie	Selten	Gelegentlich	Häufig	Fast immer
❑	❑	❑	❑	❑

13. Während der letzten 7 Tage: wie STARK war Ihre **Übelkeit** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

14. Während der letzten 7 Tage: wie HÄUFIG mussten Sie **erbrechen**?

Nie	Selten	Gelegentlich	Häufig	Fast immer
❑	❑	❑	❑	❑

15. Während der letzten 7 Tage: wie STARK war Ihr **Erbrechen** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

16. Während der letzten 7 Tage: wie STARK waren Ihre **Probleme beim Schlafen** (wie z.B. Schwierigkeiten beim Einschlafen, Durchschlafen oder zu frühes Aufwachen) im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

17. Während der letzten 7 Tage: wie SEHR haben **Probleme beim Schlafen** (wie z.B. Schwierigkeiten beim Einschlafen, Durchschlafen oder zu frühes Aufwachen) Sie in Ihren täglichen Aktivitäten GESTÖRT?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

universitätbonn

PRO-CTCAE

Interdisziplinäres Screening und Assessment zur Etablierung altersgerechter Behandlungskonzepte in der Onkogeriatrie

DIE JOHANNITER. Aus Liebe zum Leben

Patientennummer	Datum	Zyklus

18. Während der letzten 7 Tage: wie STARK war Ihre **Müdigkeit, Erschöpfung oder fehlende Energie** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

19. Während der letzten 7 Tage: wie SEHR haben **Müdigkeit, Erschöpfung oder fehlende Energie** Sie in Ihren täglichen Aktivitäten GESTÖRT?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

20. Während der letzten 7 Tage: wie STARK war Ihre **Kurzatmigkeit** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

21. Während der letzten 7 Tage: wie STARK waren **Taubheit oder Kribbeln in Händen oder Füßen** im schlimmsten Fall?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

22. Während der letzten 7 Tage: wie SEHR hat Sie **Taubheit oder Kribbeln in Händen oder Füßen** in Ihren täglichen Aktivitäten GESTÖRT?

Gar nicht	Ein wenig	Mäßig	Ziemlich	Sehr
❑	❑	❑	❑	❑

Die PRO-CTCAE Fragen wurden von der Division of Cancer Control and Population Sciences des NATIONAL CANCER INSTITUTE am NATIONAL INSTITUTE OF HEALTH in Bethesda, Maryland, U.S.A. entwickelt und sind dessen Eigentum. Es kann keine Gewähr für die PRO-CTCAE Ergebnisse übernommen werden.

Appendix D

Results of the evaluation study

Distribution of CTCAE grade ≥ 3 toxicity during therapy course per individual CARG and CRASH score items and patient characteristics

*Table D-1 Distribution of overall CTCAE grade ≥ 3 toxicity during therapy course per patient characteristics; Fisher's exact test was used for p value calculation; ECOG, Eastern Cooperative Oncology; * Fisher's exact test excluded missing values*

	No toxicity		Toxicity		
	n	**%**	**n**	**%**	**P value**
Age [years]					
70-74	4	12.1	29	87.9	0.178
75-79	12	26.7	33	73.3	
80-84	5	20.0	20	80.0	
≥ 85	0	0.0	10	100.0	
Sex					
Female	10	17.9	46	82.1	1.000
Male	11	19.3	46	80.7	
ECOG performance status					
Fully active (0)	10	26.3	28	73.7	0.239
Capable of all self-care (1-2)	10	16.7	50	83.3	
Limited or no self-care (3-4)	1	6.7	14	93.3	
Charlson Comorbidity Index					
No comorbidity (0)	10	20.8	38	79.2	0.282
Little comorbidity (1-2)	8	15.1	45	84.9	
Moderate comorbidity (3-4)	2	18.2	9	81.8	
High comorbidity (≥ 5)	1	100.0	0	0.00	
Polymedication					
No polymedication (< 5)	12	21.8	43	78.2	0.743
Polymedication (≥ 5)	7	15.9	37	84.1	
Hyperpolymedication (≥ 10)	2	14.3	12	85.7	
Tumor type					
Solid tumor	11	17.7	51	82.3	0.813
Hematological tumor	10	19.6	41	80.4	

Table D-1 (continued)

	No toxicity		Toxicity		
	n	**%**	**n**	**%**	**P value**
Cancer stage					
I	0	0.0	7	100.0	0.296*
II	3	30.0	7	70.0	
III	4	13.8	25	86.2	
IV	13	25.0	39	75.0	
Missing	1	6.7	14	93.3	
Metastasis					
No	4	16.7	20	83.3	1.000*
Yes	7	19.4	29	80.6	
Not applicable	10	19.6	41	80.4	
Missing	0	0.0	2	100.00	
Treatment intention					
Palliative	10	17.9	46	82.1	0.419*
Curative	8	17.0	39	83.0	
Chronic condition	3	37.5	5	62.5	
Missing	0	0.0	2	100.0	
Treatment type					
Chemotherapy	9	13.4	58	86.6	0.001
Targeted or immunotherapy	5	83.3	1	16.7	
Combined chemotherapy and targeted or immunotherapy	7	17.5	33	82.5	
Additional therapy					
None	13	20.3	51	79.7	0.109
Radiotherapy	2	6.9	27	93.1	
Surgery	5	35.7	9	64.3	
Radiotherapy and surgery	1	16.7	5	83.3	

*Table D-2 Distribution of overall CTCAE grade ≥ 3 toxicity during therapy course per CARG score items; GI, gastrointestinal; GU, genitourinary; *Fisher's exact test*

	No toxicity		Toxicity		
	n	**%**	**n**	**%**	**P value***
Socio-demographics					
Age [years]					
≥ 72	20	19.0	85	81.0	1.000
< 72	1	12.5	7	87.5	
Tumor/treatment variables					
Cancer type					
GI/GU tumor	2	11.1	16	88.9	0.518
Others	19	20.0	76	80.0	
Dose					
Reduced	2	12.5	14	87.5	0.732
Standard	19	19.6	78	80.4	
Number of treatment agents					
Monotherapy	3	20.0	12	80.0	1.000
Polytherapy	18	18.4	80	81.6	
Laboratory variables					
Hemoglobin [g/dL]					
≥ 10 (female), ≥ 11 (male)	18	24.7	55	75.3	0.041
< 10 (female), < 11 (male)	3	7.5	37	92.5	
Creatinine clearance Jeliffe [mL/min]					
< 34	0	0.0	11	100.0	0.213
≥ 34	21	20.6	81	79.4	
Geriatric assessment variables					
Hearing abilities					
Fair/worse	4	9.5	38	90.5	0.079
Good/excellent	17	23.9	54	76.1	

Table D-2 (continued)

	No toxicity		Toxicity		
	n	**%**	**n**	**%**	**P value***
Falls in past six months					
0	18	19.8	73	80.2	0.761
≥ 1	3	13.6	19	86.4	
Medication intake					
No assistance	21	19.1	89	80.9	1.000
Requires assistance	0	0.0	3	100.0	
Limited in walking one block					
Not limited at all	11	18.3	49	81.7	1.000
Limited	10	18.9	43	81.1	
Decreased social activity because of health/emotional problems					
A little or none of the time	18	23.1	60	76.9	0.074
Some, most, all of the time	3	8.6	32	91.4	

*Table D-3 Distribution of overall CTCAE grade ≥ 3 toxicity during therapy course per CRASH score items; IADL, instrumental activities of daily living; LDH, lactate dehydrogenase; ECOG, Eastern Cooperative Oncology Group; MMSE, Mini-Mental State Examination; MNA, Mini Nutritional Assessment; MAX2, chemotherapy toxicity index; *Fisher's exact test*

	No toxicity		Toxicity		
	n	**%**	**n**	**%**	**P value***
Hematologic score					
Diastolic blood pressure					
≤ 72	10	20.8	38	79.2	0.631
> 72	11	16.9	54	83.1	
IADL					
26-29	18	19.8	73	80.2	0.761
10-25	3	13.6	19	86.4	
LDH [U/L]					
> 0.74 x ULN	17	16.0	89	84.0	0.022
≤ 0.74 x ULN	4	57.1	3	42.9	
Nonhematologic score					
ECOG performance status					
0	10	26.3	28	73.7	0.239
1-2	10	16.7	50	83.3	
3-4	1	6.7	14	93.3	
MMS					
30	6	30.0	14	70.0	0.202
< 30	15	16.1	78	83.9	
MNA					
28-30	2	13.3	13	86.7	0.733
< 28	19	19.4	79	80.6	
All scores					
Therapy toxicity (MAX2)					
0 (0-0.44)	6	25.0	18	75.0	0.37
1 (0.45-0.57)	10	21.3	37	78.7	
2 (< 0.57)	5	11.9	37	88.1	

Incidence of severe toxicity at start of therapy (first cycle or at least 3 weeks of therapy)

Table D-4 Toxicity at start of therapy (within the first cycle or within at least the first 3 weeks if cycle was shorter) according to CTCAE ≥ 3 grade; AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma-glutamyl transferase

	Grade ≥ 3		Grade 3		Grade 4		Grade 5	
	n	%	n	%	n	%	n	%
Overall toxicity	78	69.0	74	65.5	36	31.9	3	2.7
Hematologic toxicity	57	50.4	53	46.9	27	23.9	0	0.0
Anemia	35	31.0	35	31.0	0	0.0	0	0.0
Febrile neutropenia	12	10.6	12	10.6	0	0.0	0	0.0
Neutropenia	32	28.3	12	10.6	20	17.7	0	0.0
Leukopenia	39	34.5	19	16.8	20	17.7	0	0.0
Thrombopenia	17	15.0	6	5.3	12	10.6	0	0.0
Nonhematologic toxicity	56	49.6	55	48.7	9	8.0	3	2.7
Acute coronary syndrome	2	1.8	1	0.9	0	0.0	1	0.9
Atrial flutter	0	0.0	0	0.0	0	0.0	0	0.0
Heart failure	1	0.9	1	0.9	0	0.0	0	0.0
Thromboembolic event	3	2.7	3	2.7	0	0.0	0	0.0
AST	7	6.2	6	5.3	1	0.9	0	0.0
ALT	6	5.3	5	4.4	1	0.9	0	0.0
GGT	16	14.2	12	10.6	4	3.5	0	0.0
Bilirubin	0	0.0	0	0.0	0	0.0	0	0.0
Creatinine	4	3.5	3	2.7	1	0.9	0	0.0
Proteinuria	0	0.0	0	0.0	0	0.0	0	0.0
Dyspnea	12	10.6	11	9.7	1	0.9	0	0.0
Erythroderma	0	0.0	0	0.0	0	0.0	0	0.0
Urticaria	0	0.0	0	0.0	0	0.0	0	0.0
Palmar-plantar erythrodysesthesia syndrome	0	0.0	0	0.0	0	0.0	0	0.0
Gastrointestinal bleeding	2	1.8	1	0.9	1	0.9	0	0.0

Table D-4 (continued)

	Grade ≥ 3		Grade 3		Grade 4		Grade 5	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Infections	26	23.0	21	18.6	3	2.7	2	1.8
Dehydration	2	1.8	2	1.8	0	0.0	0	0.0
Hyponatremia	9	8.0	9	8.0	0	0.0	0	0.0
Hypokalemia	9	8.0	9	8.0	0	0.0	0	0.0
Dysphagia	1	0.9	1	0.9	0	0.0	0	0.0
Dry mouth	0	0.0	0	0.0	0	0.0	0	0.0
Mucositis	1	0.9	1	0.9	0	0.0	0	0.0
Pain	13	11.5	13	11.5	0	0.0	0	0.0
Anorexia	7	6.2	7	6.2	0	0.0	0	0.0
Constipation	0	0.0	0	0.0	0	0.0	0	0.0
Diarrhea	3	2.7	3	2.7	0	0.0	0	0.0
Nausea	5	4.4	5	4.4	0	0.0	0	0.0
Vomiting	2	1.8	2	1.8	0	0.0	0	0.0
Insomnia	1	0.9	1	0.9	0	0.0	0	0.0
Fatigue	10	8.8	10	8.8	0	0.0	0	0.0
Peripheral sensory neuropathy	0	0.0	0	0.0	0	0.0	0	0.0

Relationship between the hematologic/nonhematologic CRASH score with the CARG score

*Table D-5 Relationship between the hematologic and nonhematologic CRASH score with the CARG score (n = 120); for Spearman's rho the scores were treated as continuous variables; for Fisher's exact test and weighted kappa as categorial variables; CI, confidence interval; * CRASH scores pooled into three categories by combining mid-low and mid-high categories*

Score	Spearman's rho (p value)	Weighted kappa * (CI, p value)	Fisher's exact test p value
Hematologic CRASH score vs CARG score	0.020 (0.830)	-0.001 (-0.062-0.061, 0.982)	0.680
Nonhematologic CRASH score vs CARG score	0.188 (0.039)	0.063 (-0.72-0.198, 0.333)	0.163

Relationship between the hematologic/nonhematologic CRASH score with the hematologic/nonhematologic physicians' judgments

*Table D-6 Relationship between the hematologic/nonhematologic CRASH score with the hematologic/nonhematologic physicians' judgments (n = 118); scores and physicians' judgments were treated as categorial variables; CI, confidence interval; * CRASH scores pooled into three categories by combining mid-low and mid-high categories*

Score	Spearman's rho (p value)	Weighted kappa * (CI, p value)	Fisher's exact test p value
Hematologic CRASH score vs hematologic physicians' judgments	0.094 (0.310)	-0.038 (-0.102-0.026, 0.369)	0.565
Nonhematologic CRASH score vs nonhematologic physicians' judgments	0.102 (0.271)	0.022 (-0.086-0.130, 0.668)	0.357

Comparison of the proportion of patients with hematologic toxicity per hematologic CRASH score value in our study with the original CRASH score development study

*Table D-7 Proportion of patients with hematologic severe toxicity during therapy course per hematologic CRASH score value in our study compared with the proportion of patients with hematologic toxicity per respective score value in the original CRASH development study; * from the CRASH development study* [68]

Hematologic CRASH score value	Number of patients per score value in our study	Proportion of patients with toxicity in our study [%]	Proportion of patients with toxicity in development study [%] *
0	1	0.0	7
1	4	75.0	7
2	11	36.4	23
3	27	51.9	23
4	34	76.5	54
5	32	84.4	54
6	4	50.0	100

Comparison of the proportion of patients with nonhematologic toxicity per nonhematologic CRASH score value in our study with the original CRASH score development study

Table D-8 *Proportion of patients with nonhematologic severe toxicity during therapy course per nonhematologic CRASH score value in our study compared with the proportion of patients with nonhematologic toxicity per respective score value in the original CRASH development study; * from the CRASH development study* [68]

Nonhematologic CRASH score value	Number of patients per score value in our study	Proportion of patients with toxicity in our study [%]	Proportion of patients with toxicity in development study [%] *
1	1	0.0	33
2	2	100.0	33
3	8	50.0	46
4	14	42.9	46
5	19	47.4	67
6	49	59.2	67
7	18	83.3	93
8	2	100.0	93

Kaplan-Meier estimates of the CARG score and combined CRASH score for severe overall toxicity

Table D-9 CARG score category low (n = 63): Time to occurrence of first overall severe toxicity (CTCAE grade ≥ 3) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patient	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimate	Standard error		
1	0.25	yes	0.984	0.016	1	62
2	0.50	yes			2	61
3	0.50	yes			3	60
4	0.50	yes			4	59
5	0.50	yes	0.921	0.034	5	58
6	0.75	yes	0.905	0.037	6	57
7	1.00	yes	0.889	0.040	7	56
8	1.00	no			7	55
9	1.50	yes			8	54
10	1.50	yes			9	53
11	1.50	yes			10	52
12	1.50	yes			11	51
13	1.50	yes			12	50
14	1.50	yes			13	49
15	1.50	yes	0.776	0.053	14	48
16	1.75	yes			15	47
17	1.75	yes	0.743	0.055	16	46
18	2.00	yes			17	45
19	2.00	yes			18	44
20	2.00	yes			19	43
21	2.00	yes			20	42
22	2.00	yes			21	41
23	2.000	yes			22	40
24	2.00	yes	0.630	0.061	23	39
25	2.25	yes	0.614	0.062	24	38
26	2.50	yes			25	37

Table D-9 (continued)

Patient	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimate	Standard error		
27	2.50	yes			26	36
28	2.50	yes			27	35
29	2.50	yes			28	34
30	2.50	yes	0.533	0.063	29	33
31	2.50	no			29	32
32	3.00	yes			30	31
33	3.00	yes			31	30
34	3.00	yes			32	29
35	3.00	yes	0.467	0.064	33	28
36	3.00	no			33	27
37	3.25	yes	0.449	0.064	34	26
38	3.50	yes	0.432	0.063	35	25
39	3.50	no			35	24
40	4.00	yes			36	23
41	4.00	yes	0.396	0.063	37	22
42	4.50	yes	0.378	0.063	38	21
43	5.00	yes	0.360	0.062	39	20
44	5.50	yes	0.342	0.062	40	19
45	5.50	no			40	18
46	6.00	no			40	17
47	6.75	yes	0.322	0.061	41	16
48	7.00	yes	0.302	0.061	42	15
49	8.00	no			42	14
50	8.50	yes	0.280	0.060	43	13
51	11.50	no			43	12
52	11.50	no			43	11
53	13.00	yes	0.255	0.060	44	10
54	14.00	no			44	9
55	14.00	no			44	8
56	16.50	no			44	7

Table D-9 (continued)

Patient	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimate	Standard error		
57	17.00	no			44	6
58	18.00	yes	0.212	0.063	45	5
59	19.00	yes	0.170	0.063	46	4
60	22.00	no			46	3
61	22.00	no			46	2
62	24.00	yes	0.085	0.068	47	1
63	30.50	no			47	0

Table D-10 CARG score category high (n = 50): Time to occurrence of first overall severe toxicity (CTCAE grade ≥ 3) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimate	Standard error		
1	0.50	yes			1	49
2	0.50	yes			2	48
3	0.50	yes			3	47
4	0.50	yes	0.920	0.038	4	46
5	0.75	yes	0.900	0.042	5	45
6	1.00	yes			6	44
7	1.00	yes			7	43
8	1.00	yes			8	42
9	1.00	yes	0.820	0.054	9	41
10	1.25	yes	0.800	0.057	10	40
11	1.50	yes			11	39
12	1.50	yes			12	38
13	1.50	yes			13	37
14	1.50	yes			14	36
15	1.50	yes			15	35
16	1.50	yes			16	34
17	1.50	yes			17	33
18	1.50	yes			18	32
19	1.50	yes			19	31
20	1.50	yes			20	30
21	1.50	yes			21	29
22	1.50	yes	0.560	0.070	22	28
23	1.75	yes			23	27
24	1.75	yes			24	26
25	1.75	yes	0.500	0.071	25	25
26	2.00	yes			26	24
27	2.00	yes			27	23
28	2.00	yes			28	22

Table D-10 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimate	Standard error		
29	2.00	yes			29	21
30	2.00	yes			30	20
31	2.00	yes			31	19
32	2.00	yes			32	18
33	2.00	yes			33	17
34	2.00	yes			34	16
35	2.00	yes	0.300	0.065	35	15
36	2.25	yes	0.280	0.063	36	14
37	2.50	yes			37	13
38	2.50	yes	0.240	0.060	38	12
39	3.00	yes	0.220	0.059	39	11
40	3.25	yes	0.200	0.057	40	10
41	4.50	yes			41	9
42	4.50	yes	0.160	0.052	42	8
43	5.50	no			42	7
44	5.50	no			42	6
45	5.50	no			42	5
46	7.50	yes	0.128	0.050	43	4
47	7.50	no			43	3
48	10.00	yes			44	2
49	10.00	yes	0.043	0.039	45	1
50	16.75	no			45	0

Table D-11 Combined CRASH score category low (n = 24): Time to occurrence of first overall severe toxicity (CTCAE grade ≥ 3) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
1	0.25	yes	0.958	0.041	1	23
2	0.50	yes			2	22
3	0.50	yes	0.875	0.068	3	21
4	1.50	yes			4	20
5	1.50	yes			5	19
6	1.50	yes	0.750	0.088	6	18
7	2.00	yes	0.708	0.093	7	17
8	2.25	yes			8	16
9	2.25	yes	0.625	0.099	9	15
10	3.00	yes			10	14
11	3.00	yes	0.542	0.102	11	13
12	4.00	yes	0.500	0.102	12	12
13	5.00	yes	0.458	0.102	13	11
14	5.50	no			13	10
15	5.50	no			13	9
16	6.00	no			13	8
17	6.75	yes	0.401	0.104	14	7
18	10.00	yes	0.344	0.104	15	6
19	11.50	no			15	5
20	16.50	no			15	4
21	17.00	no			15	3
22	18.00	yes	0.229	0.116	16	2
23	19.00	yes	0.115	0.100	17	1
24	24.00	yes	0.000	0.000	18	0

Table D-12 Combined CRASH score category high (n = 89): Time to occurrence of first overall severe toxicity (CTCAE grade ≥ 3) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
1	0.50	yes			1	88
2	0.50	yes			2	87
3	0.50	yes			3	86
4	0.50	yes			4	85
5	0.50	yes			5	84
6	0.50	yes	0.933	0.027	6	83
7	0.75	yes			7	82
8	0.75	yes	0.910	0.030	8	81
9	1.00	yes			9	80
10	1.00	yes			10	79
11	1.00	yes			11	78
12	1.00	yes			12	77
13	1.00	yes	0.854	0.037	13	76
14	1.00	no			13	75
15	1.25	yes	0.843	0.039	14	74
16	1.50	yes			15	73
17	1.50	yes			16	72
18	1.50	yes			17	71
19	1.50	yes			18	70
20	1.50	yes			19	69
21	1.50	yes			20	68
22	1.50	yes			21	67
23	1.50	yes			22	66
24	1.50	yes			23	65
25	1.50	yes			24	64
26	1.50	yes			25	63
27	1.50	yes			26	62
28	1.50	yes			27	61

Table D-12 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
29	1.50	yes			28	60
30	1.50	yes			29	59
31	1.50	yes	0.660	0.050	30	58
32	1.75	yes			31	57
33	1.75	yes			32	56
34	1.75	yes			33	55
35	1.75	yes			34	54
36	1.75	yes	0.603	0.052	35	53
37	2.00	yes			36	52
38	2.00	yes			37	51
39	2.00	yes			38	50
40	2.00	yes			39	49
41	2.00	yes			40	48
42	2.00	yes			41	47
43	2.00	yes			42	46
44	2.00	yes			43	45
45	2.00	yes			44	44
46	2.00	yes			45	43
47	2.00	yes			46	42
48	2.00	yes			47	41
49	2.00	yes			48	40
50	2.00	yes			49	39
51	2.00	yes			50	38
52	2.00	yes	0.421	0.053	51	37
53	2.50	yes			52	36
54	2.50	yes			53	35
55	2.50	yes			54	34
56	2.50	yes			55	33
57	2.50	yes			56	32
58	2.50	yes			57	31

Table D-12 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
59	2.50	yes	0.342	0.051	58	30
60	2.50	no			58	29
61	3.00	yes			59	28
62	3.00	yes			60	27
63	3.00	yes	0.306	0.049	61	26
64	3.00	no			61	25
65	3.25	yes			62	24
66	3.25	yes	0.282	0.048	63	23
67	3.50	yes	0.269	0.048	64	22
68	3.50	no			64	21
69	4.00	yes	0.257	0.047	65	20
70	4.50	yes			66	19
71	4.50	yes			67	18
72	4.50	yes	0.218	0.045	68	17
73	5.50	yes	0.205	0.044	69	16
74	5.50	no			69	15
75	5.50	no			69	14
76	7.00	yes	0.191	0.043	70	13
77	7.50	yes	0.176	0.042	71	12
78	7.50	no			71	11
79	8.00	no			71	10
80	8.50	yes	0.158	0.042	72	9
81	10.00	yes	0.141	0.041	73	8
82	11.50	no			73	7
83	13.00	yes	0.121	0.039	74	6
84	14.00	no			74	5
85	14.00	no			74	4
86	16.75	no			74	3
87	22.00	no			74	2

Table D-12 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
88	22.00	no			74	1
89	30.50	no			74	0

Predictive performance of the ECOG performance status and age for severe hematologic and nonhematologic toxicity

*Table D-13 Predictive performance of the commonly used predictors ECOG (Eastern Cooperative Oncology Group) performance status and age for hematologic and nonhematologic grade ≥ 3 toxicity; * Fisher's exact test*

		Calibration			Discrimination	
	Cross tables		Logistic regression		ROC curves	
	Toxicity n (%)	P value*	P value	Odds Ratio (CI)	ROC-AUC (CI)	P value
Hematologic toxicity						
ECOG						
0	23 (60.5)	0.519	0.893	1.028 (0.687-1.539)	0.526 (0.408-0.645)	0.649
1	37 (74.0)					
2	7 (70.0)					
3	9 (60.0)					
Age [years]						
70-74	26 (78.8)	0.193	0.130	0.934 (0.856-1.020)	0.390 (0.279- 0.501)	0.059
75-79	30 (66.7)					
80-84	13 (52.0)					
≥ 85	7 (70.0)					
Nonhematologic toxicity						
ECOG						
0	17 (44.7)	0.006	0.002	2.134 (1.325-3.438)	0.665 (0.566-0.765)	0.003
1	28 (56.0)					
2	9 (90.0)					
3	13 (86.7)					
Age [years]						
70-74	21 (63.6)	0.773	0.729	0.985 (0.906-1.071)	0.468 (0.361-0.575)	0.565
75-79	24 (53.3)					
80-84	16 (64.0)					
≥ 85	6 (60.0)					

Risk factors of nonhematologic and hematologic CTCAE grade ≥ 3 toxicity during therapy course

Table D-14 Univariate logistic regression for CARG and CRASH score items as well as other items to determine risk factors of hematologic CTCAE grade ≥ 3 toxicity during therapy course; categories are based on original score cut-offs for respective items; CI, confidence interval; GI, gastrointestinal; GU, genitourinary; IADL, instrumental activities of daily living; LDH, lactate dehydrogenase; ECOG, Eastern Cooperative Oncology; MMSE, Mini-Mental State Examination; MNA, Mini Nutritional Assessment; MAX2, chemotherapy toxicity index; nd: not determinable; reference: in italic; if no reference class is denoted, variables are treated as continuous

	Odds ratio (95% CI)	P value
CARG score items		
Age [years] *< 72* vs ≥ 72	1.253 (0.283-5.552)	0.767
Cancer type *Other* vs GI/GU tumor	0.549 (0.197-1.534)	0.253
Dose *Standard* vs Reduced	1.083 (0.347-3.382)	0.891
Number of treatment agents *Monotherapy* vs Polytherapy	2.719 (0.902-8.195)	0.076
Hemoglobin [g/dL] *≥ 10* vs < 10 (female), *≥ 11* vs < 11 (male)	2.636 (1.066-6.520)	0.036
Creatinine clearance Jeliffe [mL/min] *≥ 34* vs < 34	0.837 (0.229-3.061)	0.788
Hearing abilities *Good/excellent* vs Fair/worse	1.359 (0.594-3.109)	0.468
Falls in past six months *0* vs ≥ 1	0.819 (0.309-2.168)	0.687
Medication intake *No assistance* vs Requires assistance	0.233 (0.020-2.660)	0.241
Limited in walking one block *Not limited at all* vs Limited	0.652 (0.296-1.437)	0.289
Decreased social activity because of health/emotional problems *A little or none of the time* vs Some, most, all of the time	1.324 (0.555-3.156)	0.527

Table D-14 (continued)

	Odds ratio (95% CI)	P value
Hematologic CRASH score items		
Diastolic blood pressure *≤ 72* vs > 72	1.711 (0.774-3.783)	0.185
IADL *26-29* vs 10-25	0.642 (0.246-1.676)	0.365
LDH *≤ 167* vs > 167	2.949 (0.625-13.930)	0.172
Therapy toxicity (MAX2)		
0	-	
1	3.556 (1.270-9.950)	0.016
2	8.333 (2.617-26.535)	0.000
Other risk factors		
Age	0.934 (0.856-1.020)	0.130
Sex *male* vs female	1.056 (0.481-2.316)	0.893
ECOG performance status	1.028 (0.687-1.539)	0.893
Charlson Comorbidity Index	1.114 (0.790-1.571)	0.540
Creatinine Clearance (Cockcroft-Gault) [mL/min]	1.016 (0.996-1.035)	0.121
Medication number before start of cancer treatment	1.037 (0.930-1.157)	0.511
Tumor type *Solid tumors* vs Hematological tumors	1.559 (0.699-3.477)	0.278
Treatment type		
Chemotherapy	-	
Targeted or immunotherapy	0.091 (0.010-0.831)	0.034
Combined chemotherapy and targeted or immunotherapy	1.204 (0.507-2.858)	0.675

Table D-15 Univariate logistic regression for CARG and CRASH score items, and other items to determine risk factors of nonhematologic severe toxicity CTCAE grade ≥ 3 during therapy course; categories are based on original score cut-offs for respective items; CI, confidence interval; GI, gastrointestinal; GU, genitourinary; IADL, instrumental activities of daily living; LDH, lactate dehydrogenase; ECOG, Eastern Cooperative Oncology Group; MMSE, Mini-Mental State Examination; MNA, Mini Nutritional Assessment; MAX2, chemotherapy toxicity index; nd: not determinable; reference: italic; if no reference class is denoted, variables were treated as continuous

	Odds ratio (95% CI)	**P value**
CARG score items		
Age [years] *< 72* vs ≥ 72	0.462 (0.089-2.398)	0.358
Cancer type *Other* vs GI/GU tumor	4.135 (1.123-15.228)	0.033
Dose *Standard* vs Reduced	1.611 (0.520-4.993)	0.409
Number of treatment agents *Monotherapy* vs Polytherapy	1.324 (0.444-3.945)	0.615
Hemoglobin [g/dL] *≥ 10* vs < 10 (f), *≥ 11* vs < 11 (m)	2.919 (1.247-6.830)	0.014
Creatinine clearance (Jeliffe) [mL/min] *≥ 34* vs < 34	3.414 (0.702-16.598)	0.128
Hearing abilities *Good/excellent* vs Fair/worse	1.189 (0.545-2.596)	0.664
Falls in past six months *0* vs ≥ 1	1.255 (0.479-3.288)	0.644
Medication intake *No assistance* vs Requires assistance	nd	
Limited in walking one block *Not limited* at all vs Limited	1.701 (0.794-3.645)	0.172
Decreased social activity because of health/emotional problems *A little or none of the time* vs Some, most, all of the time	3.206 (1.297-7.928)	0.012

Table D-15 (continued)

	Odds ratio (95% CI)	P value
Nonhematologic CRASH score items		
ECOG performance status		
0	-	
1-2	1.987 (0.871-4.532)	0.103
3-4	8.029 (1.589-40.582)	0.012
MMS *30* vs < 30	1.583 (0.600-4.180)	0.354
MNA *28-30* vs < 28	2.473 (0.814-7.510)	0.110
Therapy toxicity (MAX2)		
0	-	
1	0.884 (0.327-2.391)	0.809
2	1.286 (0.460-3.594)	0.632
Other risk factors		
Age	0.985 (0.906-1.071)	0.729
Sex *male* vs female	2.376 (1.099-5.136)	0.028
ECOG performance status	2.134 (1.325-3.438)	0.002
Charlson Comorbidity Index	0.773 (0.559-1.069)	0.120
Creatinine Clearance (Cockcroft-Gault) [mL/min]	0.991 (0.974-1.009)	0.338
Medication number before start of cancer treatment	1.138 (1.014-1.277)	0.029
Tumor type *Solid tumors* vs Hematological tumors	0.533 (0.249-1.139)	0.104
Treatment type		
Chemotherapy	-	
Targeted or immunotherapy	0.105 (0.012-0.949)	0.045
Combined chemotherapy and targeted or immunotherapy	0.639 (0.287-1.424)	0.273

Predictive performance for the alterations of the planned treatment

*Table D-16 Predictive performance of the CARG score and combined CRASH score for total alterations of the planned treatment; * Fisher's exact test*

	Calibration				Discrimination	
	Cross tables		Logistic regression		ROC curves	
	Total alterations n (%)	**P value ***	**Wald statistic P value**	**Odds Ratio (95% CI)**	**ROC-AUC (95% CI)**	**P value**
CARG						
Low	3 (60.0)	0.327	0.086	1.125 (0.983-1.288)	0.604 (0.488-0.720)	0.077
Mid	37 (63.8)					
High	38 (76.0)					
CRASH combined score						
Low	2 (66.7)	0.380	0.107	1.200 (0.962-1.497)	0.612 (0.500-0.725)	0.057
Mid-Low	12 (57.1)					
Mid-High	47 (69.1)					
High	17 (81.0)					

*Table D-17 Predictive performance of the CARG score and combined CRASH score for minor alterations of the planned treatment (delays or dose reductions); * Fisher's exact test*

	Calibration				Discrimination	
	Cross tables		Logistic regression		ROC curves	
	All delays or reductions n (%)	P value *	Wald statistic P value	Odds Ratio (95% CI)	ROC-AUC (95% CI)	P value
CARG						
Low	3 (60.0)	0.615	0.410	0.952 (0.846-1.071)	0.462 (0.354-0.570)	0.490
Mid	24 (41.4)					
High	19 (38.0)					
CRASH combined score						
Low	2 (66.7)	0.632	0.227	0.879 (0.713-1.084)	0.446 (0.337-0.556)	0.335
Mid-Low	10 (47.6)					
Mid-High	25 (36.8)					
High	9 (42.9)					

Kaplan-Meier analysis and Cox Regression for total and minor alterations of the planned treatment

Table D-18 Kaplan-Meier analysis and Cox regression for the time until occurrence of total alterations of the planned treatment

Kaplan-Meier analysis		
	Median time to total alterations [weeks] (95% CI)	**P value (log-rank test)**
Total	5.000 (2.716-7.284)	
CARG		
Low/Mid	7.000 (2.420-11.580)	0.150
High	4.500 (3.027-5.973)	
CRASH combined score		
Low/Mid-Low	7.000 (0.904-13.096)	0.319
Mid-High/High	4.500 (2.990-6.010)	
Cox regression		
	Hazard ratio (95% CI)	**P value**
CARG Low/Mid vs High	1.371 (0.879-2.137)	0.164
CRASH combined score Low/Mid-Low vs Mid-High/High	1.330 (0.745-2.374)	0.335

Table D-19 Kaplan-Meier analysis and Cox regression for the time until occurrence of minor alterations (delays or dose reductions) of the planned treatment

Kaplan-Meier analysis		
	Median time to minor alterations [weeks] (95% CI)	**P value (log-rank test)**
Total	12.000 (6.255-17.745)	
CARG		
Low/Mid	11.500 (4.889-18.111)	0.927
High	12.000 (-)	
CRASH combined score		
Low/Mid-Low	9.000 (1.819-16.181)	0.574
Mid-High/High	12.000 (4.697-19.303)	
Cox regression		
	Hazard ratio (95% CI)	**P value**
CARG Low/Mid vs High	1.027 (0.571-1.848)	0.929
CRASH combined score Low/Mid-Low vs Mid-High/High	0.830 (0.429-1.606)	0.580

Kaplan-Meier estimates of the CARG score and the combined CRASH score for major alterations of the planned treatment (discontinuations or changes)

Table D-20 CARG score category low (n = 63): Time to occurrence of major alterations of treatment (discontinuations or changes) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
1	0.25	yes	0.984	0.016	1	62
2	0.50	no			1	61
3	0.75	yes	0.968	0.022	2	60
4	1.00	no			2	59
5	1.50	no			2	58
6	1.75	yes	0.951	0.027	3	57
7	2.00	yes			4	56
8	2.00	yes			5	55
9	2.00	yes			6	54
10	2.00	yes			7	53
11	2.00	yes	0.868	0.044	8	52
12	2.50	yes	0.851	0.046	9	51
13	2.50	no			9	50
14	2.50	no			9	49
15	3.00	no			9	48
16	3.25	no			9	47
17	3.50	yes	0.833	0.048	10	46
18	4.50	yes	0.815	0.051	11	45
19	5.50	no			11	44
20	5.50	no			11	43
21	5.50	no			11	42
22	5.50	no			11	41
23	6.00	no			11	40
24	6.00	no			11	39
25	6.50	no			11	38
26	6.75	no			11	37

Table D-20 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
27	8.00	yes	0.793	0.054	12	36
28	8.00	no			12	35
29	8.00	no			12	34
30	9.00	no			12	33
31	9.00	no			12	32
32	10.00	yes	0.768	0.057	13	31
33	11.00	no			13	30
34	11.00	no			13	29
35	11.00	no			13	28
36	11.50	no			13	27
37	11.50	no			13	26
38	11.50	no			13	25
39	11.50	no			13	24
40	11.50	no			13	23
41	12.00	no			13	22
42	12.00	no			13	21
43	12.00	no			13	20
44	12.50	no			13	19
45	13.00	no			13	18
46	13.000	no			13	17
47	14.00	yes			14	16
48	14.00	yes	0.678	0.079	15	15
49	15.50	no			15	14
50	16.50	no			15	13
51	17.00	no			15	12
52	17.00	no			15	11
53	18.00	no			15	10
54	18.50	no			15	9
55	22.00	no			15	8
56	22.00	no			15	7

Table D-20 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
57	23.00	no			15	6
58	23.00	no			15	5
59	24.00	no			15	4
60	24.00	no			15	3
61	26.00	no			15	2
62	26.50	no			15	1
63	30.50	no			15	0

Table D-21 CARG score category high (n = 50): Time to occurrence of major alterations of treatment (discontinuations or changes) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
1	0.50	yes			1	49
2	0.50	yes			2	48
3	0.50	yes	0.940	0.034	3	47
4	0.75	yes	0.920	0.038	4	46
5	1.00	yes	0.900	0.042	5	45
6	1.50	yes			6	44
7	1.50	yes	0.860	0.049	7	43
8	1.75	yes	0.840	0.052	8	42
9	2.00	yes			9	41
10	2.00	yes			10	40
11	2.00	yes			11	39
12	2.00	yes	0.760	0.060	12	38
13	2.00	no			12	37
14	4.50	no			12	36
15	4.50	no			12	35
16	5.00	yes			13	34
17	5.00	yes			14	33
18	5.00	yes	0.695	0.066	15	32
19	5.00	no			15	31
20	5.50	no			15	30
21	5.50	no			15	29
22	5.50	no			15	28
23	6.00	no			15	27
24	7.00	yes	0.669	0.068	16	26
25	7.50	yes	0.643	0.070	17	25
26	7.50	no			17	24
27	8.50	no			17	23
28	10.00	yes			18	22

Table D-21 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
29	10.00	yes	0.587	0.075	19	21
30	11.00	no			19	20
31	11.00	no			19	19
32	11.50	yes	0.557	0.077	20	18
33	12.00	yes	0.526	0.078	21	17
34	12.00	no			21	16
35	12.50	yes	0.493	0.080	22	15
36	12.50	no			22	14
37	13.00	yes	0.458	0.082	23	13
38	14.00	yes	0.422	0.083	24	12
39	14.50	yes	0.387	0.083	25	11
40	15.00	yes	0.352	0.083	26	10
41	16.75	no			26	9
42	17.00	no			26	8
43	18.50	no			26	7
44	18.5	no			26	6
45	20.00	no			26	5
46	22.50	no			26	4
47	23.00	no			26	3
48	23.00	no			26	2
49	24.50	yes	0.176	0.131	27	1
50	40.50	no			27	0

Table D-22 Combined CRASH score category low (n = 24): Time to occurrence of major alterations of treatment (discontinuations or changes) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
1	0.25	yes	0.958	0.041	1	23
2	0.50	yes	0.917	0.056	2	22
3	0.50	no			2	21
4	1.50	no			2	20
5	2.00	no			2	19
6	5.50	no			2	18
7	5.50	no			2	17
8	6.00	no			2	16
9	6.00	no			2	15
10	6.75	no			2	14
11	10.00	yes	0.851	0.082	3	13
12	11.50	no			3	12
13	11.50	no			3	11
14	12.00	yes	0.774	0.105	4	10
15	12.00	no			4	9
16	12.00	no			4	8
17	16.50	no			4	7
18	17.00	no			4	6
19	18.00	no			4	5
20	22.50	no			4	4
21	23.00	no			4	3
22	24.00	no			4	2
23	24.00	no			4	1
24	26.00	no			4	0

Table D-23 Combined CRASH score category high (n = 89): Time to occurrence of major alterations of treatment (discontinuations or changes) in Kaplan-Meier analysis; yes: severe toxicity occurred; no: censored

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
1	0.50	yes			1	88
2	0.50	yes	0.978	0.016	2	87
3	0.75	yes			3	86
4	0.75	yes	0.955	0.022	4	85
5	1.00	yes	0.944	0.024	5	84
6	1.00	no			5	83
7	1.50	yes			6	82
8	1.50	yes	0.921	0.029	7	81
9	1.75	yes			8	80
10	1.75	yes	0.898	0.032	9	79
11	2.00	yes			10	78
12	2.00	yes			11	77
13	2.00	yes			12	76
14	2.00	yes			13	75
15	2.00	yes			14	74
16	2.00	yes			15	73
17	2.00	yes			16	72
18	2.00	yes			17	71
19	2.00	yes	0.796	0.043	18	70
20	2.50	yes	0.785	0.044	19	69
21	2.50	no			19	68
22	2.50	no			19	67
23	3.00	no			19	66
24	3.25	no			19	65
25	3.50	yes	0.773	0.045	20	64
26	4.50	yes	0.760	0.046	21	63
27	4.50	no			21	62
28	4.50	no			21	61

Table D-23 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
29	5.00	yes			22	60
30	5.00	yes			23	59
31	5.00	yes	0.723	0.048	24	58
32	5.00	no			24	57
33	5.50	no			24	56
34	5.50	no			24	55
35	5.50	no			24	54
36	5.50	no			24	53
37	5.50	no			24	52
38	6.00	no			24	51
39	6.50	no			24	50
40	7.00	yes	0.709	0.049	25	49
41	7.50	yes	0.694	0.050	26	48
42	7.50	no			26	47
43	8.00	yes	0.679	0.051	27	46
44	8.00	no			27	45
45	8.00	no			27	44
46	8.50	no			27	43
47	9.00	no			27	42
48	9.00	no			27	41
49	10.00	yes			28	40
50	10.00	yes	0.646	0.054	29	39
51	11.00	no			29	38
52	11.00	no			29	37
53	11.00	no			29	36
54	11.00	no			29	35
55	11.00	no			29	34
56	11.50	yes	0.627	0.056	30	33
57	11.50	no			30	32
58	11.50	no			30	31

Table D-23 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
59	11.50	no			30	30
60	12.00	no			30	29
61	12.00	no			30	28
62	12.50	yes	0.605	0.058	31	27
63	12.50	no			31	26
64	12.50	no			31	25
65	13.00	yes	0.581	0.061	32	24
66	13.00	no			32	23
67	13.00	no			32	22
68	14.00	yes			33	21
69	14.00	yes			34	20
70	14.00	yes	0.501	0.067	35	19
71	14.50	yes	0.475	0.069	36	18
72	15.00	yes	0.449	.070	37	17
73	15.50	no			37	16
74	16.75	no			37	15
75	17.00	no			37	14
76	17.00	no			37	13
77	18.50	no			37	12
78	18.50	no			37	11
79	18.50	no			37	10
80	20.00	no			37	9
81	22.00	no			37	8
82	22.00	no			37	7
83	23.00	no			37	6
84	23.00	no			37	5
85	23.00	no			37	4
86	24.50	yes	0.337	0.110	38	3
87	26.50	no			38	2

Table D-23 (continued)

Patients	Time (weeks)	Event	Cumulative proportion of patients without toxicity		Cumulative number of events	Number of remaining patients
			Estimates	Standard error		
88	30.50	no			38	1
89	40.50	no			38	0

Appendix E

Results of the medication risk analysis

Prevalence of antineoplastic agents and supportive care medication after start of cancer therapy

Table E-1 Drug classes (ATC code level 2) and individual drugs which patients received as antineoplastic agents after start of cancer therapy (n = 128)

Drug class (ATC code level 2)	Number of drug prescriptions
Plant alkaloids and other natural products (L01C)	76
Platinum compounds (L01XA)	54
Alkylating agents (L01A)	45
Monoclonal antibodies (L01XC)	40
Antimetabolites (L01B)	30
Corticosteroids for systemic use (H02)	29
Cytotoxic antibiotics and related substances (L01D)	29
Others	8
Drug	**Number of patients**
Paclitaxel	38
Carboplatin	37
Rituximab	33
Cyclophosphamide	30
Doxorubicin	25
Vincristine	23
Prednisolone	13
Predisone	11
Oxaliplatin	10
Fluorouracil	9
Etoposide	9
Bendamustine	8
Gemcitabine	8
Cisplatin	7
Decitabine	6
Bortezomib	5
Dexamethasone	5
Methotrexate	4

Table E-2 Drug classes (ATC code level 2) and individual drugs of supportive care medication which patients received after start of therapy (n = 128)

Drug class (ATC code level 2)	**Number of drugs prescriptions**
Antiemetics and antinauseants (A04)	116
Antihistamines for systemic use (R06)	71
Corticosteroids for systemic use (H02)	62
Drugs for acid-related disorders (A02)	60
Antigout preparations (M04)	49
Analgesics (N02)	35
Detoxifying agents for antineoplastic treatment (V03AF)	34
Others	38
Drug	**Number of patients**
Ondansetron	109
Dexamethasone	62
Ranitidine	52
Clemastine	50
Allopurinol	49
Paracetamol	35
Mesna	32
Dimetindene	21
Calcium folinate	11

Distribution of CTCAE grade ≥ 3 toxicity during therapy course regarding medication risks

*Table E-3 Distribution of CTCAE grade ≥ 3 toxicity during therapy course per medication risks; PIM, potentially inadequate medication; rPDDI, relevant potential drug-drug interactions; * Fisher's exact test*

	n (%)	P value*
Overall toxicity		
Patients *without* polymedication	43 (78.2)	0.471
Patients with polymedication	49 (84.5)	
Patients *without* PIM	42 (79.2)	0.633
Patients with PIM	50 (83.3)	
Patients *without* rPDDI	60 (75.9)	0.033
Patients with rPDDI	32 (94.1)	
Hematologic toxicity		
Patients *without* polymedication	36 (65.5)	0.841
Patients with polymedication	40 (69.0)	
Patients *without* PIM	36 (67.9)	1.000
Patients with PIM	40 (66.7)	
Patients *without* rPDDI	47 (59.5)	0.008
Patients with rPDDI	29 (85.3)	
Nonhematologic toxicity		
Patients *without* polymedication	30 (54.5)	0.344
Patients with polymedication	37 (63.8)	
Patients *without* PIM	27 (50.9)	0.125
Patients with PIM	40 (66.7)	
Patients *without* rPDDI	44 (55.7)	0.298
Patients with rPDDI	23 (67.6)	