Colorectal cancer treatment and early response evaluation, how do we best evaluate treatment response?

Per Byström

Stockholm 2010
To my Family
With dedication to all patients who participated in the presented studies
ABSTRACT

Colorectal cancer (CRC) is common, being responsible for around 12% of new cancers in Sweden, and contributing heavily to the large numbers of cancer deaths yearly. CRC is the second most common cause of cancer death not only in Sweden but also worldwide.

In a large (n=567) multicentre, phase III study, the Nordic VI study, two different ways of giving 5-FU, bolus (FLv) and protracted infusion (Lv5FU2), together with irinotecan, in patients with metastatic (m)CRC was tested without any differences in progression-free survival (PFS) [9 months] or overall survival (OS) [19 months]. Fewer objective responses were seen in the FLIRI group (35% versus 49%, p = 0.001), but the metastatic resection rate did not differ (4% versus 6%, p = 0.3).

In the same study a subset (n=51) of the population was evaluated for early metabolic treatment response with [18F]-2-fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET). The mean baseline standard uptake value (SUV) for all tumor lesions per patient was higher in non-responders than in responders (mean 7.4 versus 5.6, p = 0.02). There was a strong correlation between metabolic response (changes in SUV) and objective response (p = 0.00001), with a sensitivity of 77% and a specificity of 76%. There was no significant correlation between metabolic response and PFS (p = 0.5) or OS (p = 0.1).

In the Nordic VI study, selected tumor markers at baseline and during treatment were evaluated for their ability to predict response rate (RR), PFS and OS in two sub-studies. In the first one (n=90), low levels of tissue inhibitor of metallo-protease 1 (TIMP-1) levels at baseline were correlated to a higher probability of obtaining an objective response (p =0.007). Plasma TIMP-1 scored as a continuous variable on a log scale (loge) was significantly associated with OS (p < 0.0001) and PFS (p =0.048).In the second sub-study (n=106) a significant correlation to OS was seen for baseline levels of all selected markers. In multivariate analyses with clinical parameters, TPA, CRP, SAA and TIMP-1 provided independent information. Changes during treatment, recorded as the slope gave with the exception of CA19-9 for OS less information about outcomes. The best correlation to response was seen for CEA, CA19-9 and TPA with AUC values of 0.78, 0.83 and 0.79, respectively, using a combined model based upon an interaction between the slope and the baseline value.

Health related quality of life (HRQoL) was also evaluated in a subset of the Nordic VI population (n=220). There were no differences in HRQoL between the two treatment groups at any time point. Emotional functioning and pain improved, and diarrhea worsened with time. Most baseline QoL subscales correlated with OS. Independent information on OS, but not PFS or RR was seen for physical functioning (p=.000), appetite (p=.028) and constipation (p=.041) together with hemoglobin level. A summary score, based on the sum of all scale items, was independently related to OS (p=.000) and PFS (p=.006) but not to RR. Most patients with an objective tumor response or a long (≥ 4 months) disease stabilisation had a favourable HRQoL outcome; however, a minor portion did not. No significant correlations were seen between changes in QoL parameters during treatment and RR, PFS and OS.

Keywords: Colorectal cancer, TIMP-1, tumor marker, PET, CEA, CA19-9, TPA, SAA, CRP, prealbumin, HRQoL.
LIST OF PUBLICATIONS


V. Byström P, Johansson B, Bergström I, Berglund Å, Sørbye H, Tveit KM, Glimelius B. Health-related Quality of Life as therapeutic guidance in patients with advanced colorectal cancer receiving palliative combination chemotherapy. In manuscript
Additional papers referred to in this thesis:


CONTENTS

Abstract................................................................. i
List of publications......................................................... ii
Additional papers referred to in the thesis......................... iii
Introduction................................................................... 1
Background................................................................. 1
Colorectal cancer......................................................... 1
Response evaluation in colorectal cancer......................... 6
Positron emission tomography (PET).............................. 8
Tumor markers, clinical parameters and routine blood-tests... 9
Quality of life evaluations............................................ 10
Aims of the investigations............................................ 11
Material and Methods.................................................. 12
Results.................................................................... 14
Discussion.................................................................. 20
Summary and General Conclusions................................. 25
Future perspectives..................................................... 26
Acknowledgements..................................................... 33
References............................................................... 35
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-FU</td>
<td>5-fluorouracil</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>APC</td>
<td>Adenomatous polyposis coli</td>
</tr>
<tr>
<td>BSC</td>
<td>Best supportive care</td>
</tr>
<tr>
<td>CEA</td>
<td>Carcinembryonic antigen</td>
</tr>
<tr>
<td>CRC</td>
<td>Colorectal cancer</td>
</tr>
<tr>
<td>CRT</td>
<td>Chemoradiotherapy</td>
</tr>
<tr>
<td>CT</td>
<td>Computerized tomography</td>
</tr>
<tr>
<td>DFS</td>
<td>Disease free survival</td>
</tr>
<tr>
<td>DPD</td>
<td>Dihydropyrimidine dehydrogenase</td>
</tr>
<tr>
<td>EGFR</td>
<td>Epidermal growth factor receptor</td>
</tr>
<tr>
<td>FDG</td>
<td>Fluorodeoxyglucose</td>
</tr>
<tr>
<td>GI</td>
<td>Gastro intestinal</td>
</tr>
<tr>
<td>GKM</td>
<td>Gate keeper mutation</td>
</tr>
<tr>
<td>Gy</td>
<td>Gray</td>
</tr>
<tr>
<td>HNPCC</td>
<td>Hereditary non-polyposis colorectal cancer</td>
</tr>
<tr>
<td>HR</td>
<td>Hazard ratio</td>
</tr>
<tr>
<td>HRQoL</td>
<td>Health related quality of life</td>
</tr>
<tr>
<td>LV</td>
<td>Leucovorin</td>
</tr>
<tr>
<td>MSI</td>
<td>Micro satellite instability</td>
</tr>
<tr>
<td>ONC</td>
<td>Oncogene</td>
</tr>
<tr>
<td>OS</td>
<td>Overall survival</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PFS</td>
<td>Progression free survival</td>
</tr>
<tr>
<td>QoL</td>
<td>Quality of life</td>
</tr>
<tr>
<td>RR</td>
<td>Response rate</td>
</tr>
<tr>
<td>SAA</td>
<td>Serum amyloide A</td>
</tr>
<tr>
<td>SDF-1</td>
<td>Stroma cell derived factor 1</td>
</tr>
<tr>
<td>SUV</td>
<td>Standard uptake value</td>
</tr>
<tr>
<td>TIMP</td>
<td>Tissue inhibitor of metallo-proteases</td>
</tr>
<tr>
<td>TGF-β</td>
<td>Tumor growth factor β</td>
</tr>
<tr>
<td>TP</td>
<td>Thymidine phosphorylase</td>
</tr>
<tr>
<td>TPA</td>
<td>Tissue polypeptide antigen</td>
</tr>
<tr>
<td>TS</td>
<td>Thymidylate synthase</td>
</tr>
<tr>
<td>TSG</td>
<td>Tumor suppressor gene</td>
</tr>
<tr>
<td>TTP</td>
<td>Time to progression</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
</tbody>
</table>
Introduction

Colorectal cancer (CRC) contributes heavily to the numbers of death due to cancer with CRC as one of the leading courses of death (in the developed world) and the incidence is slowly increasing. Surgery is for most CRC the only treatment that can result in permanent cure. A substantial number of patients (35-40%) will sometime over the disease period end up with inoperable disease. It is therefore of great importance to try to optimise therapy for this large group of individuals that cannot be cured. Robust and clinically useful predictive factors are still, in spite of the enormous amount of studies conducted in the area, lacking. Treatment evaluation of the effects of chemotherapy is in CRC mostly based on imaging by computerized tomography (CT) after approximately 2 months, corresponding to 3-4 cycles of treatment. Earlier evaluation of response would allow for avoidance of unnecessary toxicity and costs for non-responding patients and a change to a potentially more active treatment. Research that addresses these clinical questions is urgently needed.

Background

Colorectal cancer
The CRC incidence is approximately 6000 new cases (12% of all cancer) in Sweden (2008). Practically all cases are adenocarcinomas.

Standard treatment for patients with CRC is surgical resection of the primary and regional lymph nodes for localized disease. At diagnosis, approximately 10-15% has stage I, 30% stage II, 35-40% stage III and 20-25% stage IV. Five- year relative survival rates for the various stages is 90% (stage I), 80% (stage II), 60% (stage III) and <10% (stage IV) with presently no major difference between colon and rectum.

Adjuvant chemotherapy after colon cancer surgery reduces the risk of recurrence. 5-FU alone (modulated with calciumfolinate or given as oral capecitabine) reduces the risk by approximately 30 % and the addition of oxaliplatin reduces it further by about 20%. Since most recurrences are fatal, overall survival (OS) is also improved although not to the same extent as recurrence-free survival (1). Adjuvant treatment is routinely indicated in patients with stage III and in stage II if high-risk criteria for relapse are present. It is likely that the effects in rectal cancer are the same, although this has not been proven in large randomized trials after the introduction of better rectal cancer surgery (2).

In rectal cancer, preoperative radiotherapy, sometimes in combination with chemotherapy is routinely indicated in many patients since this therapy decreases the risk of local recurrence and possibly improves survival (2).

Treatment of Patients with Stage IV Disease:
Treatment of patients with primarily advanced or recurrent CRC depends on the location of the disease. For patients with locally recurrent and/or liver-only and/or lung-only metastatic disease, surgical resection, if feasible, is the only potentially curative treatment. At the time being only 15-20% of patients with CRC liver metastasis are candidates for resection with curative intent (3), although population-based series report lower resection rates (4). For patients with hepatic metastasis considered to be
resectable (i.e., based on limited number of lesions, intrahepatic locations of lesions, lack of major vascular involvement, absent or limited extrahepatic disease, and sufficient functional hepatic reserve), five-year survival rates of 35% to 55% have been reported (5, 6). Better surgical techniques and advances in preoperative imaging have improved patient selection for resection. In addition, studies with multiagent chemotherapy have demonstrated that patients with metastatic disease isolated to the liver, which historically would be considered unresectable, can occasionally be made resectable after the administration of chemotherapy (7).

**First-line single agent Chemotherapy:**
5-fluorouracil (5-FU) has for more than 40 years been the mainstay in the treatment of mCRC. Meta-analysis has shown that 5-FU treatment prolongs survival by about 6 months, compared to best supportive care (8). 5-FU has a short plasma half-life and is cytotoxic mainly to cells in the S-phase. Therefore, with bolus administration of 5-FU, only a small proportion of the tumor cells is susceptible as compared with continuous administration of the drug. Studies have shown different mechanisms of action of 5-FU and effects when using high-dose short-term treatment compared with long-term, low-dose exposure to 5-FU (9-11). These observations support the statement that 5-FU may be considered as two different drugs (12). A great number of different schedules with 5-FU exist, but there are traditionally two major ways of administering the 5-FU treatment, bolus injection or long term infusion (at least 24 hours) (table 1). A bolus push injection was in one study superior to a short-term infusion of 5-FU (13).

<table>
<thead>
<tr>
<th>Table 1. Some often used 5-FU schedules</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bolus injection regimens</strong></td>
</tr>
<tr>
<td>Nordic (14)</td>
</tr>
<tr>
<td>LV 5-FU day 1-2</td>
</tr>
<tr>
<td>500mg/m² 60mg/m² qw2</td>
</tr>
<tr>
<td>Bolus Bolus</td>
</tr>
<tr>
<td>Mayo (15) US</td>
</tr>
<tr>
<td>LV 5-FU day 1-5</td>
</tr>
<tr>
<td>20mg/m² 425mg/m² qw4</td>
</tr>
<tr>
<td>Bolus Bolus</td>
</tr>
<tr>
<td>Roswell park (16) US</td>
</tr>
<tr>
<td>LV 5-FU day 1</td>
</tr>
<tr>
<td>500mg/m² 600mg/m² qw1</td>
</tr>
<tr>
<td>2 h Bolus</td>
</tr>
<tr>
<td><strong>Infused regimens</strong></td>
</tr>
<tr>
<td>TTD (17), Spain</td>
</tr>
<tr>
<td>5-FU day 1-2</td>
</tr>
<tr>
<td>3500mg/m² Qw1</td>
</tr>
<tr>
<td>48 h</td>
</tr>
<tr>
<td>AIO (18) Germany</td>
</tr>
<tr>
<td>LV 5-FU day 1</td>
</tr>
<tr>
<td>500mg/m² 2600mg/m² qw1 x 7, 2w rest</td>
</tr>
<tr>
<td>2 h 24 h</td>
</tr>
<tr>
<td>Lokich (19) US</td>
</tr>
<tr>
<td>5-FU continuous</td>
</tr>
<tr>
<td>300mg/m²/d</td>
</tr>
<tr>
<td>Continuous infusion for 10 weeks or more</td>
</tr>
<tr>
<td><strong>Combination of bolus injection and infusion (hybrids)</strong></td>
</tr>
<tr>
<td>DeGramont France (20)</td>
</tr>
<tr>
<td>LV 5-FU day 1,2</td>
</tr>
<tr>
<td>200mg/m² 400mg/m² 600mg/m² qw2</td>
</tr>
<tr>
<td>2 h Bolus 22 h</td>
</tr>
</tbody>
</table>
LV=leucovorin (calcium folinate)
Several simplifications of the deGramont schedule exist.

A meta-analysis showed infused 5-FU to be superior to bolus 5-FU in terms of tumor response but not OS (21). Several studies have shown the equivalence/non-inferiority between bolus 5-FU/leucovorin and per oral 5-FU analogues, both UFT (22) and capecitabine in mCRC (23, 24). Also in the adjuvant situation after surgery for colon cancer stage III are the two drugs non-inferior to bolus 5-FU/leucovorin (25, 26).

First-line Multiagent Chemotherapy
Three randomized studies demonstrated improved response rates, progression-free survival (PFS), and OS when irinotecan or oxaliplatin was combined with 5-FU-leucovorin (27-29).

Intergroup study N9741 compared the bolus IFL schedule with the bolus/infused FOLFOX4. Patients assigned to FOLFOX4 experienced an improved PFS (median, 6.9 months vs. 8.7 months, \( P = .014 \)) and OS (15.0 months vs. 19.5 months, \( P = .001 \)) compared with patients assigned to IFL. Subsequently, two studies compared FOLFOX with FOLFIRI, and patients were allowed to cross over upon progression on first-line therapy, respectively (30, 31). PFS and OS were identical between the treatment groups in both studies. Since the publication of these studies, either FOLFOX or FOLFIRI is considered acceptable first-line treatments of patients with mCRC. The Nordic group compared in a randomized multi-centre study (paper IV) two irinotecan, 5-FU, leucovorin combinations as first line treatment, with no clear differences between schedules (32). Two of the hospitals prospectively evaluated several pre-treatment variables for their ability firstly to predict response, PFS and OS and, secondly, if early changes in these variables could predict subsequent response as well as PFS and OS. These studies are presented in the thesis as papers II, III, V and VI.

The Addition of Bevacizumab, Cetuximab and Panitumumab to Multiagent Chemotherapy:
Bevacizumab: Patients with previously untreated mCRC were randomly assigned to either IFL or IFL and the antiangiogenic antibody bevacizumab (33). PFS was significantly better (10.6 months) in the group given IFL and bevacizumab than in the group given IFL (6.2 months). OS was also significantly longer (20.3 vs 15.6 months). When bevacizumab was combined with an oxaliplatin combination (FOLFOX or Xelox), a slight but significant gain was seen in PFS (8.3 vs 7.2 months) but not in OS (22.8 vs 18.5 months) (34). In two small but randomized studies, bevacizumab also (35) improved treatment results in combination with 5-FU/leucovorin alone (OS from 14.6 to 17.9 months) (35). Similarly, bevacizumab increased response rates and prolonged PFS when combined with capecitabine and mitomycin C (from 30 to 36% and 5.7 to 8.5 months, respectively) (36). In a second-line study patients who progressed on 5-FU-leucovorin and irinotecan were randomly assigned to either FOLFOX or FOLFOX and bevacizumab. A statistically significant improvement in PFS (7.4 vs. 5.5 months) and OS (12.5 vs. 10.7) was seen (37). Based on these studies, bevacizumab can reasonably be added to either 5-FU alone, FOLFIRI or FOLFOX for patients undergoing first- or second-line treatment of mCRC, although the gains were more clear in the study using the irinotecan combination than the oxaliplatin combination. Bevacizumab in combination with FOLFOX 4 in the second line setting revealed, however, that the
hazard ratios for death demonstrated a greater magnitude of benefit than seen from the addition of bevacizumab to an oxaliplatin/fluoropyrimidine doublet in the first-line (38). For a more comprehensive overview see (39).

**Cetuximab:** This antibody against the epithelial growth factor receptor (EGFR) was first explored in the third line situation after failure on both an oxaliplatin and an irinotecan combination (40). It has later been shown that response to cetuximab can only be seen if KRAS is wild-type (41). When cetuximab recently has been explored as first-line treatment, separate analyses have therefore been performed according to KRAS mutation status, although patient inclusion was not restricted to wt KRAS tumors. KRAS mutations have then been detected in about 40% of the tumors. Median PFS for the wt population was in the OPUS study 7.7 months with the combination of cetuximab + FOLFOX and 7.2 mo with FOLFOX only but for the population with mutated KRAS the corresponding figures were 5.5 mo and 8.6 mo respectively (42). Efficacy analyses of the phase III CRYSTAL trial have shown a significant improvement in PFS, overall response, and curative surgery rate when adding cetuximab to FOLFIRI in the first-line treatment of mCRC. A statistically significant difference in favour of cetuximab was seen in KRAS wt pts for PFS p=0.0167 and best overall response 59.3% [cetuximab + FOLFIRI] vs. 43.2% [FOLFIRI], p=0.0025) (43). In contrast, the COIN study (44) evaluating the addition of cetuximab to one of two oxaliplatin-based regimens found no difference in either PFS or OS in KRAS wild-type tumors, whereas a 7% difference was seen in the response rate. The Nordic group has compared the Nordic FLOX regimen with or without cetuximab in a phase III study including 557 patients (Nordic VII). No data are presently available.

**Panitumumab:** This new fully human EGFR antibody is similar in action to cetuximab, but panitumumab (IgG2) and cetuximab (IgG1) differ in their isotype and they might differ in their mechanism of action. Similar to cetuximab, panitumumab was first explored in the third-line situation (45). Panitumumab was the first monoclonal antibody to demonstrate the use of KRAS as a predictive biomarker (46).

Two studies exploring the combination of modern chemotherapy combined with both VEGF and EGFR inhibitors were recently presented, PACCE, a randomized, phase 3 trial with either oxaliplatin- or irinotecan-based chemotherapy and bevacizumab, with or without panitumumab, as first-line treatment for mCRC (47). The results showed a significant difference in PFS (10.0 versus 11.4 months) and OS (19.4 versus 24.5) in favor of the control group in the oxaliplatin arm. In the irinotecan arm, PFS was 10.1 months for the panitumumab group and 11.7 months for the control group, with a median OS of 20.7 months for the panitumumab group and 20.5 months for the control group. The other study was, CAIRO2, a randomized, open-label, phase 3 trial that evaluated the efficacy and safety of bevacizumab and capecitabine/oxaliplatin with or without cetuximab as a first-line treatment in 755 patients with mCRC (48). Patients receiving chemotherapy, bevacizumab and cetuximab had a decreased PFS, compared with patients receiving chemotherapy and bevacizumab (9.4 versus 10.7 months), respectively.
Concomitant Chemotherapy and Radiation in Rectal Cancer:
Biologically, there is a strong case for combining drugs with radiation in a variety of solid malignancies (49, 50). The American and Norwegian experiences in rectal cancer demonstrated that postoperative CRT using fluorouracil (5-FU)-containing regimens improves both local control and OS when compared with surgery alone or with surgery followed by RT alone. (51-54). Continuous 5-FU was in one study superior to bolus in this situation (51). It has not formally been shown in randomized studies that oral drugs are equivalent to continuous infusion 5-FU during RT. Since the effects are the same when given without RT, even with a tendency to superiority for capecitabine over bolus 5-FU(23, 24) the collected international community has accepted the oral compounds (55). The bulk of clinically available documentation is presently on capecitabine (56), but also UFT have been explored (57) and used with good results (58). In rectal cancer, several randomized studies have now proven that preoperative CRT improves local control versus RT alone, although the impact on OS is still questionable (59-61). Multi-agent chemotherapy regimens combined with RT may in the future further improve the good results of neo-adjuvant CRT in rectal cancer (62, 63).

Second-line and Third-line Chemotherapy:
Second-line chemotherapy with irinotecan in patients treated with 5-FU-leucovorin as first-line therapy demonstrated improved OS when compared with infused 5-FU or supportive care (64-66). Similarly, a phase III trial randomly assigned patients who progressed on irinotecan and 5-FU-leucovorin to either infused 5-FU, oxaliplatin, or FOLFOX4. Median TTP was 4.6 months for FOLFOX4 versus 2.7 months for LV5FU2, (P < .001) (67). For patients who have progressed on irinotecan-containing regimens, a randomized phase II study was performed of either cetuximab or irinotecan and cetuximab. The median TTP for patients receiving cetuximab was 1.5 months, and the median TTP for patients receiving irinotecan and cetuximab was 4.2 months (40). Panitumumab was as described above developed for use in patients with mCRC refractory to chemotherapy (45).

The availability of several drugs with activity in mCRC has meant that survival has improved substantially and median survival has reached 20-24 months in the most recently reported trials (33, 43). This could be compared with a median survival of 6 - 8 months in the trials performed in the 1980s. Patient selection, better imaging detecting metastatic disease at an earlier stage contribute to the marked difference, but the cytostatic drugs and most recently the biologic agents are mainly responsible. The patients included in the trials are however highly selected, and typically less than 10% of eligible patients are included. At a population level, median survival in patients with metastatic disease is in the order of 11 months (68). In patients who initiated chemotherapy, it was about 15 months. For the patients suitable for treatment the importance of receiving all available drugs has been illustrated by the results presented by Grothey et al. as seen by figure 1 below.
Response-evaluation in colo-rectal cancer

All anti-cancer treatments need to be continuously evaluated to assure benefit for the patient and cost-effectiveness for the clinic. The relation between tumor response and survival is an important issue for patients with advanced CRC. The shrinkage of measurable metastatic lesions has for long been the cornerstone in the development of cytotoxic therapies (69). In CRC, as in many other solid tumor forms, the main evaluation is done by assessing structural changes, mainly by radiological evaluation and mostly based on imaging by computerized tomography (CT) after approximately 2 months, corresponding to 3-4 cycles of treatment. The RECIST criteria (70) is generally accepted as standard for CT-evaluation of chemotherapy and has recently been updated (71). The new criteria are as follows for target lesions;
Table 2. Updated RECIST criteria, version 1.1

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Response (CR)</td>
<td>Disappearance of all target lesions. Any pathological lymph node (whether target or non-target) must have reduction in short axis to &lt;10 mm</td>
</tr>
<tr>
<td>Partial Response (PR)</td>
<td>At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.</td>
</tr>
<tr>
<td>Progressive Disease (PD):</td>
<td>At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).</td>
</tr>
<tr>
<td>Stable Disease (SD):</td>
<td>Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.</td>
</tr>
</tbody>
</table>

Meta-analyses from clinical trials in advanced CRC patients have revealed that an increase in tumor response rate translates into an increase in overall survival for the patients. However, in the context of individual trials, benefits on tumor response do not allow accurate prediction of the ultimate benefit on survival (72). It has been surprisingly difficult to show a correlation between an objective response and a survival benefit (73). The reasons for this are multifold. Patients who achieve a response must have lived at least until the response was evaluated, although this guarantee-time can be compensated for by using the land-mark method (74). Patients who achieve a response are generally better patients with less tumor burden and a naturally longer survival than non-responding patients. This can only partly be compensated for in multivariate analyses since many of the factors influencing survival is only partly known and not always registered in the trials.

Patients with mCRC who achieve a CR to systemic combination chemotherapy either alone or with multimodality approach reveal as suspected a survival benefit (figure 1) (75). A CR is however seldom achieved. Using modulated 5-FU treatment, it was early shown that not only patients who achieve a CR or a PR but also those who achieve disease stabilization for at least 4 months (SD 4) have their lives prolonged (76).
Earlier evaluation of response would allow for avoidance of unnecessary toxicity and costs for non-responding patients and a change to a potentially more active treatment. Imaging of tumor metabolism is the basis for current efforts to provide such an early evaluation. The relevance of such a strategy will be discussed.

Several studies have also been conducted for evaluation of various tumor markers, to try to find reliable ways to evaluate treatment effects by other means than by assessing structural changes.

**Positron Emission Tomography**

Positron emission tomography (PET) using F-18-fluorodeoxyglucose (FDG) can visualize the enhanced glucose utilization in tumor tissue. The amount of FDG accumulated is proportional to the rate of glucose utilization (77). The Standardised Uptake Value (SUV) was introduced as a semi-quantitative measure of $^{18}$F-FDG uptake (78). Previous studies on the reproducibility of the FDG signal in malignant tumors indicate that PET imaging can reliably measure changes by more than 20% of the baseline value (79). A change of SUV by more than 20% in the tumor during treatment is accepted as a surrogate for tumor response, in accordance with the EORTC guidelines (80).

Several studies indicate that reduction of tumor glucose uptake after chemo- or radiotherapy correlates with tumor regression and patient outcome in different tumor types (81). Changes in glucose metabolism may also allow prediction of subsequent response before the reduction of tumor size (81-87).

In CRC patients, the experience is so far limited and chiefly restricted to small studies in patients with liver metastases (88, 89, 90, 91). One report showed that $^{18}$F-FDG PET 4-5 weeks after start of chemotherapy predicted the effect with an overall sensitivity of 100% and specificity of 90% on a per-lesion basis in 18 patients (91). In another, somewhat larger study in 50 heterogeneously treated patients, the FDG-PET changes after 2 months predicted PFS and OS (89). Previous studies have also indicated that FDG-PET is of value in the diagnostic work-up of patients with colorectal liver metastases (92-94).
Tumor markers, clinical parameters and routine blood-tests

Several studies have identified prognostic factors for survival in mCRC. A large pivotal multivariate analysis of 3825 mCRC patients treated with FA-FU, identified four key biomedical parameters: performance status, white blood cell (WBC) count, alkaline phosphatase and the number of involved metastatic sites (95). That this index is prognostic also in patients treated with combinations of chemotherapy (96). The identification of additional independent or even more powerful prognostic factors could have important implications for routine clinical practice and research for many reasons, including a helpful guide in treatment decision-making. The numerical amount of studies on blood markers for detection, prognosis, prediction and treatment evaluation in CRC makes it difficult to give a complete overview of the topic. The following text therefore gives a brief summary of the literature with special focus on treatment evaluation with the guidance of some recently published review articles (97-99).

In the ASCO guidelines (97) the value of the most commonly used tumor markers are described:

1. Carcinoembryonic antigen (CEA) is not recommended as a screening test for CRC. As prognostic tool the use of CEA renders support by several studies (100-102), specifically a study of 2,230 patients demonstrated that preoperative CEA was an important independent prognostic variable in predicting outcome (103) and a study of 1,146 rectal patients confirmed that preoperative CEA level was a highly significant prognostic covariate besides stage and grade (104). CEA is considered to be the marker of choice for monitoring metastatic CRC during systemic therapy. An elevation within the first weeks following chemotherapy should be interpreted with caution especially after oxaliplatin use (105, 106).

2. CA 19-9 is not recommended for use in CRC due to lack of supporting data. In pancreatic cancer the data to support the use of CA 19-9 for monitoring therapy is insufficient but it is recognised that CA 19-9 can be measured at the start of the treatment and if there is an elevation of serial CA19-9 determinations, this may be an indication of progressive disease (107, 108). Later studies have not supported the use of serial CA19-9 measurement (109).

3. DNA ploidy and Flow Cytometric Proliferation Analysis (% S-phase) are not recommended for prognostic information in early stage CRC.

4. p53 expression or mutation are insufficient for screening, diagnosis, staging, surveillance or monitoring treatment of patients with CRC. A large comprehensive systematic review (110) of 168 reports in a total of 18,766 patients concluded that with current methods of assessment, p53 status is a poor guide to both prognosis and response or resistance to therapy in patients with CRC.

5. RAS is not recommended to use for screening, diagnosis, staging, surveillance or monitoring treatment due to conflicting results. Several reported studies show Kras mutation is an adverse prognostic indicator, but the studies have wide variability in their specific results (111, 112). A recent study encompassing 1564 patients in stages II and III could not detect any prognostic information from Kras mutation status (113). The predictive role of ras is complicated by the variety of chemotherapeutic agents and regimens used, but, as described above, it is fundamental in the evaluation prior to treatment with EGFR-inhibitors.

6. Thymidylate Synthase (TS), Dihydropyrimidine Dehydrogenase (DPD) and Thymidine Phosphorylase (TP) are not recommended for prognosis of CRC and the
evidence are insufficient to recommend the use of them as predictors of response or for monitoring response to therapy.

*TS* is the rate limiting step in the biosynthesis of thymidine, one of the four nucleotides required for DNA synthesis and cell proliferation. TS is blocked by a metabolite (FdUMP) of 5-fluorouracil but is also a target for capecitabine and tegafur. In a meta-analysis of 32 published studies of TS expression with various techniques (IHC; RT-PCR or enzyme assays), the conclusion was that overall survival but not disease free survival was poorer with high TS expression in both the advanced and the adjuvant settings (114). Furthermore, it appeared that in the adjuvant surgery group, the association between expression of high intra-tumoral TS and poorer survival was strongest in those patients treated by surgery alone and the association was markedly decreased in patients who received adjuvant therapy. Recent studies suggest that polymorphisms in the promoter and the untranslated regions of the gene may be associated with different levels of TS protein in tumor (115), and may also be associated with prognosis and response to FU-based chemotherapy (116-118).

*DPD* is the major enzyme that catabolises 5-FU, DPD converts 5-FU to fluoro-5,6-dihydrouracil (FUH₂) in a rate-limiting step. More than 80% of the of the catabolism occurs in the liver were the majority of DPD is concentrated (119). DPD is mainly important in predicting toxicity and few data are available on the role of sequential values of DPD in patients undergoing therapy. Data support the thesis that inhibiting DPD increases 5-FU efficacy as in treatment with either tegafur or eniluracil (120).

7. **Microsatellite Instability** (MSI) is present in 10-15% of all patients with CRC. Seventeen series have addressed MSI and prognosis and eleven of them found a correlation between MSI high tumors and better survival compared to MSI low / MSI stable tumors (97). There are conflicting results in the literature, whether MSI status is predictive for treatment response or not. Some data suggest that MSI status might predict efficacy of adjuvant FU chemotherapy. A recent study reported that MSI predicted improved response to adjuvant therapy with an irinotecan combination (121).

**Quality of life evaluations**

Health-related quality of life (HRQoL) surveys have been used during several decades for the evaluation of QoL in cancer trials (122). Patient-reported outcomes (PROs) have become standard in oncology trials and contribute both to decision making and provide prognostic information in addition to well established clinical parameters in advanced cancers in general (123-127), as well as in mCRC (128-133). While the number of cancer clinical trials including HRQOL assessment is increasing, there is also evidence that the potentially invaluable insights that HRQOL data provide into the treatment and care of patients may not be adequately reported (134-136). For example, in a meta-analysis of randomized, controlled trials on the use of palliative chemotherapy for incurable (recurrent and distant metastatic) head and neck cancer, it was found that reports of all studies included survival and tumor response but did not include outcomes that would measure palliative benefits. And in spite of that almost 25% of patients with head and neck cancer ranked cure as only the second or third most important outcome, with items related to symptoms and functional outcomes, including appearance, ranking among the top three most important (137). A minimum set of
criteria for assessing the reported outcomes in cancer clinical trials is necessary, to compare treatments, and to make informed decisions in clinical practice. In a study, evaluating HRQoL data in a set of prostate cancer trials, a checklist developed for this was presented, and it was found that HRQoL is a valuable source of information in RCTs of treatment in metastatic prostate cancer (138).

It has further been shown that baseline measurements of HRQoL give additional predictive information for patients with advanced cancer of other diagnoses (126, 132, 139, 140). The prognostic value of changes in HRQoL scores during treatment has also been examined, however, with mixed results (141, 142). Evaluations of randomized clinical trials with QoL measurements as part of the objectives have displayed a significant learning curve in HRQoL trial reporting since the early 1990s, and also that the quality of such HRQoL reports has improved over time (143). A meta-analysis of individual patient data from EORTC clinical trials has also revealed that HRQOL data can help to predict survival in patients with cancer (144). It can therefore be expected that HRQOL data will increasingly impact on clinical decision making and treatment policies in the near future.

Aims of the investigations

The overall aim of these investigations was to increase our understanding of how to monitor palliative chemotherapy.

The specific questions addressed in the studies were the following:

- Does a more convenient bolus 5-FU regimen with irinotecan give the same PFS as a bolus-infused schedule with irinotecan. Secondary end-points in this study were OS, RR and toxicity.

- Is it possible to predict treatment response for patients with mCRC with FDG-PET early on in the course of treatment, is there a correlation between an early metabolic response and structural changes evaluated with CT as reference, and does metabolic response correlate with OS and PFS/TTP?

- Is there a correlation between high levels of TIMP-1 and resistance to chemotherapy in patients with mCRC?

- Can serum and plasma tumor markers predict patient outcome to palliative chemotherapy before the start of treatment, and can changes in these markers during the course of treatment predict outcome evaluated on a group level and for the individual patient?
Can base-line HRQoL evaluations and early changes in HRQoL among patients with mCRC, treated with combination chemotherapy, give prognostic and predictive information about RR, OS, and PFS/TTP?

Material and Methods

Patients

Paper I
A Nordic randomized, multi-center, phase III study conducted between 2001 and 2004. The study included 567 patients with histologically confirmed CRC adenocarcinoma and non-resectable metastatic disease comparing irinotecan in combination with either the Nordic bolus 5-FU and folinic acid schedule (FLv) or the bolus/infused de Gramont schedule (Lv5FU2). No prior chemotherapy other than adjuvant 5-FU-based chemotherapy completed at least 6 months before the study entry was allowed.

Papers II, III, IV and V
These four studies were conducted on subpopulations of the NORDIC VI study (see above). Patients in some of the centres were asked if they accepted participation in the studies described in paper II, III, IV and V. In the PET-study (paper II) quantitative FDG-PET was performed before treatment and after the second cycle of chemotherapy in a subset of 51 patients from two centers (Stockholm, Uppsala) that accepted the additional two PET examinations. In the TIMP-1 study (paper III) ninety patients from two centers (Stockholm, Uppsala) were included. Plasma TIMP-1 and serum CEA were measured in samples obtained before the first cycle of chemotherapy and after 2, 4 and 6 weeks of treatment. In the tumor marker study (paper IV) approximately 106 patients from three centres, Stockholm, Uppsala and Malmö, Sweden were evaluable for consecutive tumor marker measurements before start of treatment and after 2, 4 and 8 weeks of treatment. In the HRQoL study (paper V) patients at four centres, Stockholm and Uppsala, Sweden and Oslo and Bergen, Norway, were asked to participate in the QoL sub-study, altogether 220 patients accepted participation.

All studies were approved by the local Ethics committee and the clinical studies were also approved by the Swedish Medical Products Agency (Läkemedelsverket).
Methods

Papers II, III, IV and V are all parts of the Nordic VI study, paper I, where patients with advanced CRC, received irinotecan in combination with 5-FU. In Europe this combination usually is given according to the Lv5FU2 schedule (or variants of this hybrid regimen) [irinotecan 180 mg/m(2) on day 1, FA 200 mg/m(2), 5-FU bolus 400 mg/m(2) and infused 5-FU 600 mg/m(2) on day 1 and 2 (Lv5FU2-IRI)]. In the Nordic countries on the other hand it is given as irinotecan with the Nordic 5-fluorouracil (5-FU) and folinic acid (FA) bolus schedule [irinotecan 180 mg/m(2) on day 1, 5-FU 500 mg/m(2) and FA 60 mg/m(2) on day 1 and 2 (FLIRI)]. The Nordic VI study evaluated efficacy and safety of these two schedules.

In paper II, PET examinations were made 1–14 days before start of treatment and immediately before the third cycle. Examinations were made after at least 4-h fasting and testing that the patient was normoglycemic. Series of consecutive 10-min scans including the trunk and the neck was initiated 60 min after i.v. administration of 400 MBq FDG. Image analysis and assessment of respons were carried out by a nuclear medicine physician blinded for the clinical and radiological evaluation. Each scan was read visually in iterative and filtered back projected reconstructions. For the standard uptake value (SUV), a semiautomatic approach was applied using the MultiModality software by Hermes Medical Solutions (Stockholm, Sweden). Up to five index lesions were chosen in the liver and/or the lungs in the pre-treatment study. The SUVs were calculated in regions of interest (ROI) drawn manually around the target lesions. A change of SUV by >25% during treatment was accepted as a surrogate for tumor response, according to the European Organization for Research and Treatment of Cancer guidelines. A qualitative, visual PET response assessment, before and independent from the SUV calculations and radiological evaluation, was done. In half of the cases, a similar blinded evaluation was done by another nuclear medicine physician with complete agreement in all cases.

In paper III, analyses of TIMP-1 and CEA were performed, EDTA plasma samples were drawn before start of chemotherapy. The samples were stored at –80°C until analyzed. In addition, in 75 of the patients, plasma samples were collected at 2, 4, and 6 weeks of treatment (samples were drawn before the initiation of each cycle). Plasma levels of TIMP-1 were determined by use of an established, validated TIMP-1 ELISA (145).

In paper V, serum and plasma for marker analyses were taken at baseline within one week prior to the first treatment cycle and immediately (1 – 3 days) prior to cycles 2, 3 and 5, i.e. at 4 times during the first 2 months of treatment. Immediately after sampling, aliquots were frozen at -20°C for later analyses.

In paper V, the HRQoL was assessed by the Quality of Life Questionnaire C30 version 3 (QLQ-C30), developed and validated by the European Organization for Research and Treatment of Cancer (EORTC) (146). The 30-item questionnaire incorporates five functional scales and global health status (scored after transformation of the scores from 100 to zero with 100 for perfect functioning) and several symptom scores (scored zero to 100, with zero for no symptoms). The instrument was completed by the patients.
immediately prior to randomization and after 2 and 4 months prior to the tumor evaluations after 4 and 8 cycles. Missing values were handled as recommended by the EORTC manual (147). For an overall assessment of HRQoL we used not only the global health status/QoL scale (global QoL) of QLQ-C30, as recommended by EORTC (146), but also a score based on the sum of all items in the scale(148). From the sum of the functional scales and global QoL score, where 100 is the best score, the sum of the 3 symptom scales and 5 of the single items, where zero is the best score, was subtracted. The Summary QoL scale scores could range from a minimum of -800 to a maximum of +600. A relevant change in global QoL score, in selected symptom scores and in the Summary QoL scale was based on at least half a standard deviation difference from baseline, as recommended to be of clinical relevance (149-151). Depending upon the size and direction of any change, a patient could receive either a rating of ‘favourable’, ‘unfavourable’, or ‘unchanged’ QoL at each time point.

In paper V, serum and plasma for marker analyses were taken at baseline within one week prior to the first treatment cycle and immediately (1–3 days) prior to cycles 2, 3 and 5, i.e. at 4 times during the first 2 months of treatment. Immediately after sampling, aliquots were frozen at -20°C for later analyses.

**Results**

**Paper I**
Patient characteristics were well balanced. PFS did not differ between groups (median 9 months, P = 0.22). OS was also similar (median 19 months, P = 0.9)(figure 4). Fewer objective responses were seen in the FLIRI group (35% versus 49%, P = 0.001) but the metastatic resection rate did not differ (4% versus 6%, P = 0.3). Grade 3/4 neutropenia (11% versus 5%, P = 0.01) and grade 2 alopecia (18% versus 9%, P = 0.002) were more common in the FLIRI group. The 60-day mortality was 2.4% versus 2.1%.
Figure 3. Kaplan–Meier curves of (A) progression-free and (B) overall survival.

Paper II
On the 51 assessable patients SUV evaluation was carried out on median five tumor lesions per patient (1–8). The majority of tumor lesions were located in the liver [involved organs; liver ($n = 48$), lungs ($n = 13$) and lymph nodes ($n = 11$)]. The mean baseline SUV for all tumor lesions per patient was higher in non-responders than in responders (mean $7.4$ versus $5.6$, $P = 0.02$). There was a strong correlation between metabolic response (changes in SUV) and objective response (table 4).
Table 3. Correlation between metabolic response in $[^{18}\text{F}]-2$-fluoro-2-deoxy-D-glucose positron emission tomography and subsequent best overall response according to RECIST

<table>
<thead>
<tr>
<th>OBJECTIVE RESPONSE</th>
<th>PET RESPONSE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR</td>
<td>PR</td>
</tr>
<tr>
<td>CR</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>PR</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>SD</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>PD</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>22</td>
</tr>
</tbody>
</table>

Spearman rank order correlation $P$ level = 0.00007, sensitivity = 0.77 and specificity = 0.76. (PET, positron emission tomography; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease)

There was no significant correlation between metabolic response and TTP ($P = 0.5$) or OS ($P = 0.1$) figure 4

**Figure 4.** Kaplan–Meier curves comparing the proportion of surviving patients over time (months) for

(a) positron emission tomography (PET) responders ($n = 24$) and PET nonresponders ($n = 27$) ($P = 0.11$) and

(b) for objective responders ($n = 22$) and objective nonresponders ($n = 29$) ($P = 0.0002$).
Paper III
Analysis of best objective response (CR or PR vs SD and PD) showed that patients with low plasma TIMP-1 had higher probability of obtaining an objective response \( [P = 0.007] \). CEA treated as a continuous variable was also a statistically significant predictor of no response \( (P = 0.02, \text{area under the curve} 0.66) \) but much less so. Plasma TIMP-1 was the only significant covariate in a multivariable analysis of best objective response \( (P = 0.001) \). Plasma TIMP-1 scored as a continuous variable on the log scale \((\log_{e})\) was significantly associated with OS \( (\text{OS} < 0.0001) \) and with TTP \( (P = 0.048) \). Multivariable analysis showed that plasma TIMP-1 was significant for OS when including routine clinical baseline covariates \( (P < 0.0001) \). A multivariable analysis including TTP instead of OS showed that only plasma TIMP-1 was retained in the model \( \text{HR}, 1.5 \). CEA was not significantly associated with TTP or OS when TIMP-1 was included in the model.

![Kaplan-Meier curves showing the association between plasma TIMP-1 values and OS.](image)

Fig. 5. Kaplan-Meier curves showing the association between plasma TIMP-1 values and OS. The number of events (disease related deaths) and number of patients at risk at different time points during the observation period for each of the two groups are given below the figure. Solid line, patients with plasma TIMP-1 levels below the median; broken line, patients with plasma TIMP-1 levels above the median.

Paper IV
A significant correlation to OS was seen for baseline levels of all markers. Independent information was provided by TPA, CRP, SAA and TIMP-1. The baseline values of CEA, TPA and TIMP-1 were also significantly correlated to PFS and TPA to RR. Changes during treatment, i.e. the slope gave with the exception of CA19-9 for OS less information about outcomes. The best correlation to response was seen for CEA, CA19-9 and VPA with AUC values of 0.78, 0.83 and 0.79, respectively, using a combined model based upon an interaction between the slope and the baseline value.
Table 4. The area under the curve (AUC) measures the ability of each marker to correctly classify RR, OS and PFS for the three different methods

<table>
<thead>
<tr>
<th></th>
<th>CEA AUC</th>
<th>CA19-9 AUC</th>
<th>TPA AUC</th>
<th>CRPH AUC</th>
<th>SAA AUC</th>
<th>Trans AUC</th>
<th>TIMP AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=91</td>
<td>n=90</td>
<td>n=90</td>
<td>n=88</td>
<td>n=87</td>
<td>n=88</td>
<td>n=101</td>
</tr>
<tr>
<td>Intercept, Slope, Interaction</td>
<td>0.78</td>
<td>0.83</td>
<td>0.79</td>
<td>0.72</td>
<td>0.72</td>
<td>0.63</td>
<td>0.68</td>
</tr>
<tr>
<td>Slope</td>
<td>0.70</td>
<td>0.77</td>
<td>0.53</td>
<td>0.54</td>
<td>0.53</td>
<td>0.52</td>
<td>0.53</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.66</td>
<td>0.57</td>
<td>0.72</td>
<td>0.63</td>
<td>0.66</td>
<td>0.61</td>
<td>0.62</td>
</tr>
<tr>
<td>AUC OS (19 months after inclusion)</td>
<td>n=80</td>
<td>n=79</td>
<td>n=79</td>
<td>n=77</td>
<td>n=76</td>
<td>n=77</td>
<td>n=83</td>
</tr>
<tr>
<td>Intercept, Slope, Interaction</td>
<td>0.71</td>
<td>0.78</td>
<td>0.86</td>
<td>0.77</td>
<td>0.75</td>
<td>0.79</td>
<td>0.85</td>
</tr>
<tr>
<td>Slope</td>
<td>0.66</td>
<td>0.65</td>
<td>0.51</td>
<td>0.56</td>
<td>0.49</td>
<td>0.56</td>
<td>0.50</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.58</td>
<td>0.63</td>
<td>0.80</td>
<td>0.69</td>
<td>0.67</td>
<td>0.72</td>
<td>0.78</td>
</tr>
<tr>
<td>AUC PFS (9 months after inclusion)</td>
<td>n=88</td>
<td>n=87</td>
<td>n=87</td>
<td>n=85</td>
<td>n=84</td>
<td>n=85</td>
<td>n=98</td>
</tr>
<tr>
<td>Intercept, Slope, Interaction</td>
<td>0.72</td>
<td>0.71</td>
<td>0.70</td>
<td>0.62</td>
<td>0.54</td>
<td>0.63</td>
<td>0.67</td>
</tr>
<tr>
<td>Slope</td>
<td>0.54</td>
<td>0.60</td>
<td>0.57</td>
<td>0.51</td>
<td>0.50</td>
<td>0.56</td>
<td>0.54</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.67</td>
<td>0.56</td>
<td>0.71</td>
<td>0.58</td>
<td>0.56</td>
<td>0.61</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Paper V
There were no differences in HRQoL between the two treatment groups at any time point. Emotional functioning and pain improved, and diarrhea worsened with time. Most baseline QoL subscales correlated with OS. Independent information on OS, but not PFS or RR was seen for physical functioning (p=.000), appetite (p=.028) and constipation (p=.041) together with hemoglobin level. The summary score was independently related to OS (p=.000) and PFS (p=.006) but not to RR. Most patients with an objective tumor response or a long (≥ 4 months) disease stabilisation had a favourable HRQoL outcome, however, a minor portion did not. No significant correlations were seen between changes in QoL parameters during treatment and RR, PFS and OS.
Table 5. Predictive value for OS, PFS and RR of baseline clinical and QoL characteristics in stepwise multivariate analyses, statistically significant values shadowed.

<table>
<thead>
<tr>
<th>Overall Survival n=200</th>
<th>PFS n=199</th>
<th>RR n=188</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p  HR 95% CI lower</td>
<td>95% CI upper</td>
</tr>
<tr>
<td></td>
<td>p  HR 95% CI lower</td>
<td>95% CI upper</td>
</tr>
<tr>
<td></td>
<td>p  HR 95% CI lower</td>
<td>95% CI upper</td>
</tr>
</tbody>
</table>

**Step 1**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.586 1.006 .983 1.030</td>
<td>.311 .844 .607 1.172</td>
<td>.748 .993 .954 1.034</td>
</tr>
<tr>
<td>Sex</td>
<td>.465 .880 .624 1.240</td>
<td>.257 .988 .968 1.009</td>
<td>.905 .962 .508 1.821</td>
</tr>
</tbody>
</table>

**Step 2**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin per unit</td>
<td>.007 .983 .970 .995</td>
<td>.007 .984 .972 .996</td>
<td>.663 .995 .973 1.017</td>
</tr>
<tr>
<td>No organs involved</td>
<td>.019 1.475 1.065 2.044</td>
<td>.287 1.187 .866 1.627</td>
<td>.011 2.201 1.199 4.041</td>
</tr>
<tr>
<td>Alcaline Phosphatase</td>
<td>.544 1.000 .999 1.001</td>
<td>.747 1.000 .999 1.001</td>
<td>.182 .999 .997 1.000</td>
</tr>
</tbody>
</table>

**Step 3 Alternative 1 QoL baseline**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Haemoglobin per unit</td>
<td>.013 .984 .971 .997</td>
<td>.007 .984 .972 .996</td>
<td>.974 1.000 .977 1.023</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>.000 .970 .954 .986</td>
<td>.077 .988 .974 1.001</td>
<td>.577 .993 .969 1.018</td>
</tr>
<tr>
<td>Role functioning</td>
<td>.0518 1.003 .994 1.012</td>
<td>.510 .997 .990 1.005</td>
<td></td>
</tr>
<tr>
<td>Emotional functioning</td>
<td>.736 1.002 .992 1.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social functioning</td>
<td>.682 1.002 .993 1.011</td>
<td>.155 .994 .986 1.002</td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td>.688 1.003 .989 1.017</td>
<td>.093 .989 .976 1.002</td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>.930 1.000 .991 1.009</td>
<td>.117 1.007 .998 1.016</td>
<td>.655 .996 .980 1.012</td>
</tr>
<tr>
<td>Dyspnoea</td>
<td>.317 .996 .987 1.004</td>
<td></td>
<td>.807 1.002 .987 1.017</td>
</tr>
<tr>
<td>Insomnia</td>
<td>.856 1.001 .994 1.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appetite</td>
<td>.028 1.009 1.001 1.017</td>
<td>.255 1.005 .996 1.014</td>
<td>.174 1.012 .995 1.029</td>
</tr>
<tr>
<td>Nausea, vomiting</td>
<td></td>
<td>.763 .998 .983 1.013</td>
<td>.142 1.024 .992 1.056</td>
</tr>
<tr>
<td>Constipation</td>
<td>.041 1.008 1.000 1.015</td>
<td>.534</td>
<td>.995 1.010</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>.242 1.005 .997 1.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global health status</td>
<td>.993 1.000 .987 1.013</td>
<td>.977</td>
<td>.987 1.013</td>
</tr>
</tbody>
</table>

**Step 3 Alternative 2 QoL summary score baseline**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Haemoglobin per unit</td>
<td>.018 .985 .972 .997</td>
<td>.011 .985 .973 .997</td>
<td>.806 .997 .975 1.020</td>
</tr>
<tr>
<td>Summary QoL</td>
<td>.000 .998 .997 .999</td>
<td>.006 .999 .998 1.000</td>
<td>.220 .999 .997 1.001</td>
</tr>
</tbody>
</table>
Discussion

Treatment guidance and response evaluation for patients with advanced Colo-Rectal Cancer:
Much research has focused on finding better predictors of tumor response and survival, and a number of factors have been identified (97). None of presently available factors or algorithms has sufficient discriminative ability to work on an individual level. A predictive biomarker or a set of biomarkers robust enough to be useful in the clinic is therefore urgently needed. A reliable and early response evaluation is also entirely necessary to avoid inefficient treatment and avoid unnecessary toxicity for the individual patient. This was the rationale for the tumor markers explored and for the evaluation of early metabolic response described below.

Metabolic response evaluation
Our study presented in paper II showed in line with most other published studies a good correlation between an early metabolic response and later best objective response. Most researchers have also stopped the reporting of their results here, fully satisfied with a highly statistically significant result. We asked ourselves whether this should be used clinically to guide treatment decisions. Our conclusion then is not in favour of using FDG-PET for response evaluation in the clinical routine setting, since the additional information gained does not contribute sufficiently considering the extra costs and time consumed. Further the early metabolic response evaluation with PET did not give us information about long time outcome. Our study is compared to the other published studies regarding response evaluation on patients with mCRC, relatively large, prospective with a homogenous study population treated with a uniform therapy. The number of patients not responding was however limited. A similar study in the second-line situation, when fewer patients respond could reach another conclusion. Like several others before us we found that a low SUV value at baseline predicted treatment response. In one study this also fell out in a survival benefit (152).

The reason for the inferior long time prediction by PET could be explained by several factors. Firstly, an early metabolic response could in a patient with a tumor with a fast up-regulation of resistance mechanisms lead to a treatment refractory disease relatively early in the course of treatment. Secondly, a tumor with a high proliferation index could initially respond on treatment with a subsequent decreased metabolic activity, but due to fast repopulation over time still proliferate. Thirdly, when it comes to response evaluation of chemotherapy it should also be noted that earlier studies have shown that previous chemotherapy lowers sensitivity of FDG-PET when restaging patients after neo-adjuvant chemotherapy before liver surgery. Akhurst et al (153) stated that sensitivity of FDG-PET in the detection of CRC metastases during preoperative staging was decreased in patients pre-treated by neo-adjuvant chemotherapy due to down regulation of hexokinase activity (lesion detection sensitivity: 63% vs 77%). No lesions larger than 1.2 cm were missed in the untreated group, but lesions up to 3.2 cm were missed after neo-adjuvant chemotherapy. So interpretation of FDG-PET data should be done with caution in the context of concomitant chemotherapy. It has also been shown that depending on the drug and regimen, cytotoxic treatment has a clear influence on CRC cell lines. Oxaliplatin, 5FU and irinotecan cause decreased FDG uptake after 72 h.
due to decrease in glucose transport, depending on a decrease in hexokinase activity (154). Findley et al (91) were the first to report the effect of 5FU on liver metastases with FDG-PET, and found that patients with a later response on PET had a “flare” phenomenon with an increased tumor to normal liver ratio 1-2 weeks after treatment. Finally the histo-pathology of the tumor affects the reliability of the PET examinations, in mucinous CRC tumors PET underestimated the extent of tumor burden (155). With all this in respect it seems reasonable to assume that both the treatment population, the specific cytostatic agent(s) prescribed and the timing of the PET scanning after treatment might be of relevance for the interpretation of studies regarding FDG-PET as a response evaluation tool. It also appears impossible to give one single definition of metabolic response, since cut-off values depend on type of treatment, timing of evaluation and tumor type.

In conclusion, further research is needed.

TIMP-1
As most types of conventional chemotherapy, among other ways, kill cancer cells by inducing apoptosis, it could be hypothesized that a marker of apoptosis protection would be useful as marker of chemotherapy resistance, an approach evaluated in paper III. The combination of irinotecan and 5-FU has been shown to induce apoptosis in human colon cancer cell lines (156). An in vivo study using a human tumor xenograft model demonstrated that administration of CPT-11 before 5-FU resulted in higher cure rates and sensitivity to this combination was associated to induction of apoptosis (157). TIMP-1 inhibits, in addition to its matrix metalloproteinase inhibitory function, apoptosis (158-161). It has been shown that plasma levels of TIMP-1 in patients with primary CRC are associated with patient outcome; i.e., high plasma TIMP-1 predicts shorter patient survival (162-164). One explanation to this association is that TIMP-1 protects cancer cells against the apoptotic stimuli that constitutively affect the cells. In support of an anti-apoptotic function of TIMP-1 it has been shown that in a cohort of patients with metastatic breast cancer, high tumor tissue levels of TIMP-1 were significantly associated with resistance to chemotherapy (165). The study presented in paper III shows that high plasma levels of TIMP-1 in patients with mCRC treated with first-line combination chemotherapy are indicative of low probability of objective response to chemotherapy. This decreased probability of obtaining an objective response to chemotherapy in plasma high TIMP-1 patients was reflected in a significantly decreased TTP and OS of these patients. Based on the present results, it is, however, not possible to clearly separate the effect of TIMP-1 on prognosis or on objective response to chemotherapy. On the individual level TIMP-1 is not recommended to use for prediction of treatment response due to the low sensitivity and specificity with an area under the curve (AUC) of 0.67 in ROC analysis.

Other Tumor Markers
In paper IV we found that in addition to several clinical/biomedical parameters, as reported by others(95), all tested baseline tumor marker levels were significantly correlated to OS. Prognostic information has also been reported for several of the single markers in various tumor types (166-170). Whether the baseline level of either of the markers should have an impact on the decision to start palliative chemotherapy or
choose between regimens have been much debated (171, 172). The tumor marker CEA is widely used in different settings in spite of some obvious limitations. Approximately 10-15% of patients with mCRC do not have increased values and there is a daily variation of CEA serum levels by 25-30% (173). In our study the probability to respond objectively and clinically were relatively high with markedly elevated (above 10 x UNL) levels and even with values 1000 times the UNL, objective responses were seen for individual patients. Given the knowledge about the positive effects of palliative chemotherapy (27-29) on OS and QoL, and the relative weakness for CEA and the other markers to predict treatment response, a strategy using tumor marker baseline levels to select patients for treatment could be questioned. Regarding the selection of a specific treatment based on baseline marker levels, there are other more selective molecular markers more directly related to a particular drug, likely needed (172).

The only established marker to use in addition to structural imaging for monitoring of CRC treatment is CEA (97). In our study the predictive value of CEA was among the three highest with an AUC of 0.78 with a modest sensitivity and better specificity, 45% and 90%, respectively, for an optimal decrease during treatment. Since the changes in CEA levels during treatment were insignificantly correlated to both OS and TTP, the predictive strength for a treatment decision in an individual patient could be questioned. The value of CEA is well established as a prognostic marker, as well as for surveillance after surgery but to our knowledge, no large prospective studies supporting the ASCO recommendation (97), that “CEA is the marker of choice for monitoring metastatic colorectal cancer during systemic therapy” exists, and the recommendation may be questioned. That conclusion is however in contradiction to the results presented in a recently published study in mCRC (174). In their study, using the same methodology as us, they found that response defined by a CEA slope less than -0.2, corresponding to a half-life of 3.4 days, a half life rarely seen in our study population, resulted in an optimal sensitivity of 75% and a specificity of 83% with an AUC value of 0.85. Their study is somewhat larger, 51 responders compared to 42 in ours. They measured CEA levels at the six first cycles, i.e. for one month longer than we did. It is possible that the clinical value of measuring the slope increases with time. In the Iwanicki-Caron et al study, responses are divided into four groups without any information about the size of the groups and their cutoff level for the highest overall accuracy is a slope of < -0.2 over 12 weeks, a reduction in CEA only seen in a minor proportion of our responders over 8 weeks of treatment. It is, in our opinion, therefore difficult to evaluate the value of the additional information gained compared to the structural imaging (CT), the clinical standard after 8 weeks of treatment.

CA19-9 was the only marker in our study where the changes during treatment were significantly associated with OS and borderline to PFS. The ASCO guidelines (97) state that there is no place for CA19-9 for monitoring response in mCRC. Thus, it is reasonable to conclude that early changes are not valuable for prediction of either whether a response will be seen of survival.

Regarding the acute phase reactants (CRP, SAA and TTR) they behaved as a group different compared to TPA and TIMP-1, who gained much of their prognostic strength by the baseline value, and the more established markers CEA and CA19-9, who responded more to treatment with a larger difference between the AUC values at baseline and the values where the slopes were taken into consideration. The acute phase reactants behaved somewhat in between and our interpretation is that the inflammation response to the tumor is mainly dependent on the tumor phenotype and the
immunogenic response, i.e. the ability for the immune system to define the tumor as non-self, and to a lesser extent to the alteration of the tumor burden as a reaction to treatment response.

In conclusion, baseline tumor markers and selected clinical parameters provide prognostic information about survival in patients with mCRC. The ability of the individual tumor markers to predict treatment response may be relevant for certain patients, in line with clinical experience, but is for most patients not sufficient.

Health Related Quality of Life

The differences in HRQoL between the two treatment arms during the first four months of treatment were small and non-significant (paper V). These results add to the conclusions of the entire trial (paper I), namely that a bolus 5FU-irinotecan schedule (FLIRI) is not inferior to an infused schedule (Lv5FU2-Iri) as palliative treatment (the same PFS, about 9 months, OS about 19 months and similar toxicity, with the exception of neutropenia and alopecia which were seen slightly more often with FLIRI), and now also similar HRQoL. The study also found that HRQoL before start of therapy gives additional prognostic information about OS and PFS beyond what is accomplished by well-established clinical factors (95). Since we did not see any correlation between HRQoL parameters and RR, baseline HRQoL tells us more about the individual patient and his/her disease burden and the ability to tolerate treatment than about the biology of the tumor.

However, the predictive value for RR, PFS and OS of HRQoL changes during treatment was in this study insignificant, as also reported in a study in advanced breast cancer (142). The lack of predictive capability by the HRQoL changes may be related to many factors, but the relatively large group of non-responders that achieved favorable (35%) or unchanged (25%) QoL during treatment may have contributed.

The majority of the patients who clinically responded to the treatment (CR+PR and SD4, these groups behaved similarly) had either a favourable or an unchanged HRQoL, although a not insignificant proportion of the patients reported deterioration of their HRQoL. Our data indicates that this was mostly due to treatment toxicity. It was early reported that patients with mCRC who clinically responded to FA/FU had, in addition to prolonged OS (76), an improved or stabilised HRQoL (175). But it is less clear whether response to combination treatment also results in improved QoL or if the additional toxicity hampers the QoL, thus being detrimental for the palliative patient. In the three phase III trials comparing FU/FA± irinotecan or oxaliplatin (27-29), QoL scores were similar in the groups despite significant differences in RR, PFS and toxicity (i.e. neutropenia, nausea, vomiting and diarrhea). In the two trials combining 5-FU with irinotecan, the time to deterioration of global QoL was significantly prolonged in the combination arm. This latter way of reporting QoL outcome is however influenced by survival time. The reasons for the discrepancy between higher RRs and prolonged PFS on the one hand, and stable QoL scores may be multi-fold. It can in part be explained by poor compliance in follow-up assessments, leading to a potential selection bias [selective attrition] (176). Compliance was fairly low (e.g. 59-62% in the Douillard trial), and patients with more severe toxicity and/or poorer QoL may not have completed the questionnaires to the same extent as those with less severe toxicity or
better QoL. The results may thus not be representative of the whole population. Other reasons may be insufficient sensitivity of the EORTC QLQ-C30, and changes in patients’ internal standard due to a response-shift with an adaptation to the limitations due to disease or treatment (177).

It is, however, also possible that QoL evaluations do not correspond to treatment toxicity and treatment outcome as well as anticipated. In two overviews, of randomized clinical trials of palliative chemotherapy in patients with advanced CRC (178, 179) it was concluded that QoL and toxicity are different aspects that do not necessarily correlate with each other. In our study the overall finding was the same. However, we found an unfavourable QoL after 16 weeks of treatment, in spite of treatment response, in 15% using the global QoL scale and in 26% using the Summary QoL scale. This could mainly be explained by treatment related toxicity counterbalancing the improvements caused by the anti-tumor effects with more fatigue (25 out of 30), nausea/vomiting (20 out of 30) and diarrhea (20 out of 30) than before start of treatment. Gastrointestinal toxicity is frequent using combinations of 5-FU and irinotecan. More unfavourable outcomes are to be expected using the Summary QoL scale than the global QoL scale. The global QoL scale is the patient’s own evaluation of their life quality, and thus the summary measure to be used. It is also the one recommended by EORTC. In our previous studies evaluating the benefits of palliative chemotherapy in gastrointestinal cancer, we noticed that the global QoL scale was comparably insensitive to changes caused by the treatments. Although conceptually wrong, we explored several alternative ways of evaluating QoL during palliative chemotherapy in diseases with short survival times. The simple and robust Summary QoL scale turned out to be more sensitive to changes (148). It was therefore added as an additional measure to better understand the balance between favourable effects on tumor progression by chemotherapy and its toxicity.

Another finding was that a favourable or stable HRQoL was also frequently seen in the group of patients without a clinical response. Again, the patients completing the questionnaires may not be representative for the entire group. Some of the patients with only short-lived SD or PD at the first evaluation may have been in a very good shape with few if any symptoms from the disease at baseline and after 2 months of treatment. More patients responding to the questionnaire had an unfavourable HRQoL outcome in the SD2+PD group than in the CR+PR+SD4 group. This difference was more evident using the Summary QoL scale than using the global QoL scale. The toxicity to the irinotecan combination had a greater impact on the HRQoL responses from the patients where it was not counterbalanced by the favourable impact on tumor burden. Had more patients in the SD2+PD group responded to the questionnaire, it is likely that the differences would have been even more marked. Finally, SD, using RECIST criteria, may mean both up to a 30% decrease in the sum of the tumor diameters and an increase up to 20%.

Earlier studies by us have shown that patients classified as SD4 frequently have a decrease in tumor size, although not qualifying as PR and those with SD2 an increase, not qualifying as PD after 2 months (175). Symptomatic patients who respond to chemotherapy have less severe symptoms and better physical functioning compared with non-responders (129, 175, 180). Non-responders have significantly higher depression, pain and physical symptom scores compared with responders (129). In asymptomatic patients, response to 5FU/FA ± oxaliplatin chemotherapy does not
Significantly influence the patients QoL (27). In another trial, patients with good QoL scores at baseline were more likely to deteriorate than patients with poor QoL score at baseline, however, the latter were more likely to have improved scores (181).

Summary and general conclusions

1. As a palliative treatment, a simpler bolus 5-FU regimen with irinotecan (FLIRI) can be an alternative to a bolus-infused schedule (Lv5FU2-IRI) since it results in the same PFS, OS (and HRQoL, see below p 6) without being substantially more toxic. It, however, leads to slightly fewer objective responses. In special circumstances, this may be important. Both schedules are routinely used for many mCRC patients, although a slight simplification of the Lv5FU2-IRI schedule, FOLFIRI, is preferred.

2. The additional value of FDG-PET for response evaluation is for patients with CRC not supported by the study, neither in isolation nor in combination with other still limited experience. However, PET performance will probably improve by the ongoing rapid technical development. More studies are needed to define the clinical role of PET imaging for treatment assessment in mCRC. The use of PET as a staging tool is, however, well established and the additional information gained compared to conventional staging with structural imaging is often useful in the selection of treatment for the individual patient.

3. Baseline plasma TIMP-1 levels are significantly and independently associated with objective response, TTP, and OS in patients with mCRC receiving combination chemotherapy. The sensitivity and specificity of changes during treatment is, for the individual patient however too low, to support the use of TIMP-1 as marker to decide whether to continue treatment or not.

4. Although baseline tumor marker levels of selected markers, CEA, CA19-9, TPA, CRP, SAA, TTR and TIMP-1 are significantly associated to OS and to a lesser extent to RR and PFS, their sensitivity and specificity to predict treatment response for the individual patient is too low to justify an implementation in the routine assessment of treatment response for patients with mCRC. For selected patients the use of tumor markers can give additional response information as a complement to structural imaging.

5. HRQoL can give information about QoL changes during treatment and the toxicity of selected treatment regimens, but the information has to be interpreted in the context of the uneven distribution of response to the HRQoL examination, with a much higher grade of responders in the group achieving tumor treatment response. The routine use of HRQoL instruments for response evaluation is not supported by the current literature.

6. An evaluation of changes in tumor size, as e.g. measured using RECIST criteria on CT images, is superior to (early) changes in FDG-PET uptake, tumor marker levels and HRQoL.
Future perspectives

The papers constituting the basis for this thesis deal mainly with treatment response prediction and evaluation in patients with advanced CRC. The results of these prognostic and predictive papers are both positive and negative, in respect to the selected methods and markers evaluated, but the lack of usefulness in a clinical setting is somewhat of a disappointment. An evaluation of changes in tumor size, as e.g. measured using RECIST criteria on CT images, is superior to (early) changes in FDG-PET uptake, tumor marker levels and HRQoL. Our findings are not altogether unexpected, given the complexity of the task in the light of the almost exponentially growing knowledge of tumor biology, since the studies were designed. Let us briefly look upon some of the obstacles in finding robust and clinically useful predictive factors and start with some theoretical difficulties.

The cancer stem cell concept: There are several experiments that have cast serious doubts on the notion that all cells within a neoplastic cell clone are biologically equivalent. Three different experiments taking advantage of the ability to separate living cancer cells by fluorescence-activated cell sorting (FACS) have supported the concept of a small population of self-renewing, tumorigenic cells and large populations of more differentiated cells that have little, if any, ability to proliferate in vivo. In the first experiment, the FACS technique enabled researchers to segregate populations of acute myelogenous leukemia (AML) cells into majority and minority populations; the latter representing less than 1% of the neoplastic cells. As few as 5000 cells in the minority subpopulation were able to produce new tumors upon injection into host mice, in contrast to as many as 500 000 AML cells from the majority subpopulation that were unable to seed a tumor (182-185). The cells in the majority subpopulation exhibited many of the attributes of differentiated cells and had limited ability to proliferate. In another experiment with breast cancer cells prepared directly from tumors, the minority tumorigenic cell population represented only about 2% of the neoplastic population and only 200 of these cells formed a tumor in host mice, while as many as 20 000 cells from the majority cell population failed to do so. Both subpopulations contained equivalent proportions of cells in the active growth cycle (186). The third experiment was conducted with brain tumor stem cells with a similar result, as described above, when injected in mice (187). In CRC the origin of tumor formation is thought to be stem cells in the colonic crypt and with defective APC functioning β-catenin levels remain high and proliferating, still undifferentiated cells (purple in figure 6 below) fail to migrate upward and accumulate within crypts and ultimately generate adenomatous polyps. The stem cell is increasingly recognized as the target of initiating events in cancer formation (188) and there are indications that epigenetic disruption of stem/progenitor cells is a key determinant not only of cancer risk and formation but of tumor progression and heterogeneity late in the course of the tumors that arise from these cells (189).
It is therefore likely that the different subpopulations described above have different genetic and epigenetic expressions and respond different to treatment, which makes it hard to select representative markers for prediction of treatment effect and to select the appropriate treatment based only on evaluation of unselected tumor material or blood samples. Especially if most of the genetic and epigenetic changes occur in the pool of progenitor cells as suggested in the model below, see figure 7.
According to this model, cancer arises in three steps. First is an epigenetic alteration of stem/progenitor cells within a given tissue, which is mediated by aberrant regulation of tumor-progenitor genes (TPG). This alteration can be due to events within the stem cells themselves, the influence of the stromal compartment, or environmental damage or injury. Second is a gatekeeper mutation (GKM) (tumor-suppressor gene (TSG) in solid tumors, and rearrangement of oncogene (ONC) in leukemia and lymphoma). Although these GKMs are themselves monoclonal, the expanded or altered progenitor compartment increases the risk of cancer when such a mutation occurs and the frequency of subsequent primary tumors (shown as separately arising tumors). Third is genetic and epigenetic instability, which leads to increased tumor evolution. Note that many of the properties of advanced tumors (invasion, metastasis and drug resistance) are inherent properties of the progenitor cells that give rise to the primary tumor and do not require other mutations (highlighting the importance of epigenetic factors in tumor progression).

The epigenetic changes in cancer:
Cancer is widely perceived as a heterogeneous group of disorders with markedly different biological properties, which are caused by a series of clonally selected genetic changes in key tumor-suppressor genes and oncogenes. However, a growing bulk of data suggests that cancer has a fundamentally common basis that is grounded in a polyclonal epigenetic disruption of stem/progenitor cells, mediated by 'tumor-progenitor genes'. Furthermore, tumor cell heterogeneity is due in part to epigenetic variation in progenitor cells, and epigenetic plasticity together with genetic lesions drive tumor progression (189). The main epigenetic alterations seen in cancer are hypomethylation of DNA and hypoacetylation of chromatin, as well as gene-specific
hypomethylation and hypermethylation (190). Global hypomethylation, universally seen in CRC, generally arises earlier and are strongly linked to chromosomal instability and increased tumor frequency which have been shown in mouse models (191, 192) and also to loss of imprinting (193-195). In addition, the silencing of tumor-suppressor genes is associated with promoter DNA hypermethylation and chromatin hypoacetylation (196). Loss of imprinting (LOI) of the insulin growth factor 2 gene (IGF2) is a common epigenetic variant in adults and associated with a 5-fold increased frequency of colorectal neoplasia (193). Studies of DNA methylation in tumor tissue have revealed at least as many epigenetic as genetic alterations for a given gene. Presently, CRC is more and more divided into at least five subgroups depending upon MSI-status and methylator status (197).

This knowledge about epigenetic changes and the genetic instability and plasticity makes it understandable why it is hard to predict treatment outcome and select treatment strategy based on a genetic fingerprint from the tumor at the time of the primary operation. It also stresses the importance of a thorough evaluation of the tumor phenotype before start of therapy and also highlights the importance of evaluation of drug resistance factors during the course of treatment.

The role of the tumor stroma and the conductive microenvironment

Tissue architecture is critical for cell homeostasis and tissue-specific functions (198) and disruption of tissue structure usually parallels the loss of tissue-specific differentiation, suggesting that tissue architecture is intimately linked to function (199, 200). Normal epithelial cells in culture and in vivo have a defined apical-basal polarity, which is established by cell extra-cellular matrix (ECM) and cell-cell adhesions and which contributes to induction and maintenance of tissue specificity (199). Cytoplasmic molecules are also asymmetrically localized in polarized epithelial cells, for example, PIP3 and PI3K, key integrators of signaling events downstream of integrins and receptor tyrosine kinases, localized predominantly to the basal surface of polarized acinar structures in 3D cultures (201). And the transcriptional pattern varies for the same cell type depending on the environment (202). ECM remodeling enzymes such as MMPs are able to modulate the tissue architecture within the context of normal organ development and biology, and forced expression of MMPs can lead tumorogenesis in vivo (203-205).

In a manifest tumor, the tumor cells and their stroma co-evolve. Growth factors and chemokine production by fibroblasts and immune cells is altered, leading to direct stimulation of tumor cell growth and recruitment of precursor cells, which themselves respond with abnormal growth and proliferation. Malformed tumor vessels contribute to tumor hypoxia, acidosis and increased interstitial fluid pressure, an environment supporting hypomethylation and genetic instability. The tumor in turn responds with a unique repertoire of gene expression and epigenetic changes, which in turn acts to alter cell growth, invasion and ultimately metastasis. Fibroblasts are the main cellular component of tumor stroma, these cancer-associated fibroblasts (CAFs) are functionally and phenotypically distinct from normal fibroblasts that are in the same tissue but not in the tumor environment.
CAFs are identified by their spindloid appearance and the expression of α-SMA; characteristics shared by activated fibroblasts in wounds. Activated fibroblasts can initiate cancer at divergent sites including stomach and prostate (206, 207). Among factors released in the tumor stroma that have been shown to impact tumor behavior are TGF-β, SDF-1, MMPs, TIMP-1, VEGF and HIF-1-α. TGF-β is a growth inhibitor and potent immunosuppressive factor and normally contributes in regulating proliferation and apoptosis. Cancer cells either lose the growth inhibitory response to TGF-β or usurp the pathway with a stimulatory response on the cancer cells; in fact many tumors gain the ability to express TGF-β which then acts in an autocrine fashion. Additionally, TGF-β induces the epithelial mesenchymal transition (EMT) in cancer cells, aiding local invasion as well as facilitating metastatic spread (208).

SDF-1 expression is up-regulated in inflamed tissue (209), in wounds and in cancer, where it attracts cells expressing the receptor CXCR4. CXCR4 is expressed by an array of cancer cells, and the receptor ligand interaction functions as a mitogen for tumor cells and induces migration in a gradient-specific fashion resulting in local invasion as well as enabling cells to metastasize to distant sites expressing SDF-1 such as the bone marrow and peripheral organs (see the part below about the metastatic niche). Marrow-derived endothelial progenitor cells also chemotax to a SDF-1 gradient, and are recruited to tumor sites for neovascularization (210). To facilitate the restructuring needed for neovascularization, fibroblasts, macrophages and endothelial cells express and secrete MMPs via a complex tumor-stroma crosstalk (211).

MMPs act to hydrolyze the extracellular proteins of the surrounding tissue which include collagen, laminin, elastin, fibrinogen, fibronectin and vitronectin (203, 212). MMPs also have additional target proteins which include other proteinases, proteinase inhibitors, clotting factors, chemokines and chemotactic factors, growth factors, a variety of cell surface receptors and cell matrix adhesion molecules (213). MMPs have also been implicated in initiating the EMT and in promoting genomic instability (204), affording them a prominent role in tumor progression. TIMP-1, a marker evaluated in two of the presented papers in this thesis, has been shown to have an anti-apoptotic effect on bone marrow stromal cells through the PI3-kinase and JNK signaling pathways independent of the effect of TIMP on MMP activity (214). These findings suggest an effect of TIMP on cells within the tumor environment which are of marrow origin including bone marrow derived tumor fibroblasts, bone-marrow derived endothelial cells within tumor vasculature and bone-marrow derived tumor cells. VEGF is secreted by CAFs and is implicated in angiogenesis, ECM remodeling, generation of inflammatory cytokines and hematopoetic stem cell development. It also plays a crucial role in recruiting VEGF-R1 positive hematopoietic bone marrow progenitor cells to peripheral organs to initiate the pre-metastatic niche (215). Hypoxia correlates with aggressive behavior of tumors and resistance to therapy. Hypoxia-inducible factors (HIFs) are cellular transcription factors involved in the response to environmental stress. Important targets of the HIF system which are relevant to cancer biology include MDR-1 (216) IGF-2 (217), telomerase (218), and CXCR4/SDF-1 (219).
Figure 8.
This figure depicts the pre-metastatic, micrometastatic to macrometastatic transition. 

a | In response to growth factors secreted by the primary tumor, including vascular endothelial growth factor A (VEGFA), placental growth factor (PlGF) and transforming growth factor (TGF), inflammatory S100 chemokines and serum amyloid A3 (SAA3) are upregulated in pre-metastatic sites leading to clustering of bone marrow-derived haematopoietic progenitor cells (HPCs). Platelet-deployed stromal-derived growth factor 1 (SDF1) is also chemotactic for C-X-C chemokine receptor 4 (CXCR4)-positive HPCs and metastatic tumor cells (MTCs). HPCs secrete a variety of pre-metastatic factors including tumor necrosis factor-α (TNF), matrix metalloproteinase 9 (MMP9) and TGF. Activated fibroblasts, possibly derived from mesenchymal stem cells (MSCs), secrete fibronectin, an important adhesion protein in the niche, and lysyl oxidase (LOX) expression is increased, modifying the local extracellular matrix.

b | MTCs engraft the niche to populate micrometastases. The site-specific expression of adhesion integrins on activated endothelial cells such as P-selectin and E-selectin may enhance MTC adhesion and extravasation at these sites, and cell–cell interactions such as CD44 ligation in the metastatic niche may promote MTC survival and enable proliferation.

c | Recruitment of endothelial progenitor cells (EPCs) to the early metastatic niche mediates the angiogenic switch and enables progression to macrometastases. (220)

All this emerging knowledge about the tumor stroma and the role of the microenvironment makes it understandable why an examination of only the tumor cells
or secreted factors related solely to the tumor cells is insufficient to select treatment strategy and predict treatment response.

In this survey of theoretical problems with the selection of treatment and the selection of predictive markers we have not touched upon all the practical problems associated with the task. We all know that today there is a shortage of representative samples, both from the primary tumor and the surrounding microenvironment, as well as from metastasis, partly due to the lack of well established bio-banks. Our knowledge about factors that govern the response and toxicity on a given treatment for the individual patient is gradually evolving, but validated and clinically available tests are still lacking. And for the vast majority of the group of patients participating in the studies described in this thesis, the prospect of cure is, at least for the time being, not available. This fact emphasizes the importance of also looking on the QoL for our patients, bearing in mind that some of the techniques used for extracting information about the tumor is invasive and potentially hazardous and unpleasant.

To justify routine tests on our patients with the aim to individualize treatment there should also be a possibility to use the information gathered for the benefit of our patient, i.e. in the present situation with a lack of targeted therapies directed towards all the molecular events underlying the disease, there is a gap between our knowledge and our ability to transform that knowledge into therapeutic interventions. That fact should not hinder us from gathering knowledge about these molecular events, but in a structured way in the form of clinical trials. It should also encourage us to demand the incorporation of predictive and prognostic biomarkers in future clinical trials.

What strategy should we then apply to improve the selection of treatment and predictive markers, for the benefit of our patients with CRC and other GI cancers?

Even without insight into the molecular origins of human cancer, it has become increasingly clear that the traditional ways of classifying cancers have limited utility. Truly useful diagnoses must inform the clinician about the underlying nature of the disease and, more important, how each disease entity will respond to various types of therapy. The rapid technical development with different array techniques, with the gene expression arrays being the first has potentially made it possible for clinicians to stratify cancer in subgroups with distinct biological properties and prognosis, the use of bioinformatics have made it possible to identify a subset of genes whose expression correlates with a specific biological phenotype, drug responsiveness and/or prognosis.

Beyond these genes expression analysis stands a generation of novel diagnostic tools involving proteomics, in which the spectrum of proteins expressed in a patient’s tumor or serum provide critical information. The long term goal of all new analytic techniques is to individualize the treatment for each patient, based on the expression, or lack of expression, of relevant tumor and host parameters, important for the diagnosis and prognosis of the disease and for the selection of proper treatments and the selection of predictive markers to monitor the effect of the chosen therapy. Some of this information, related to tolerance/effect of the more established drugs used for the treatment of CRC patients is known, for irinotecan (221, 222), 5-FU (223) and oxaliplatin (117). Established is also the selection of patients for EGFR inhibitor therapy (43, 224). But both due to the present lack of knowledge and the extent of the topic, it would, in this thesis lead too far, to try to describe an optimal investigation. Collaboration around the difficult task is urgently needed for the benefit of our patients.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who have directly or indirectly contributed to the completion of this work, Special thanks to:

Bengt Glimelius, my supervisor, whose profound qualifications, both scientific and personal, have been crucial for the realization of this thesis and for my approach to clinical and research work. His incomparable knowledge, generous encouragement and support have been endless.

Åke Berglund, my co-supervisor, for friendship and for all scientific and clinical knowledge you generously shared with me, for your positive and inspiring attitude, for lessens in floor ball, discussions late nights over a glass of wine and good laughter’s.

Jan-Erik Frödin, head of the GI-section, for clinical guidance and support in all possible ways, both in my work with the thesis and in my clinical everyday practice.

Ulrike Garske, for your expertise and all time and effort you spent on the PET examinations and for good friendship.

Hans Jacobsson, for professional help with the PET examinations and for valuable support with the thesis.

Anders Sundin, for expertise help with the radiological evaluations and valuable support with the thesis

Peter Nygren, for intellectual stringent support and guidance in the writing of the thesis.

Birgitta Johansson, for your expertise and all good guidance in the field of HRQoL, and all warm and friendly support.

Anders Larsson, for all help with the laboratory work and for good guidance with the thesis.

Ida Bergström, for all your effort and excellent work with the statistical preparation of the QoL evaluations.

Lisa Wernroth, UCR, for all the good ideas and competence in the statistical field and for your guidance into terra incognita.

Inger Hjertström-Östh, for your professional and enthusiastic secretarial help, and for your warm and helpful manner.

Peter Ragnhammar, David Edler and Marja Hallström for good collaboration and for creating a stimulating research atmosphere.
All the staff at the research units at Karolinska and Uppsala, for all competent assistance in the clinical trials, especially Anna-Lena Johansson, Marie Rosén and Lena Bernrup.

All the colleagues who participated in the Nordic VI study, especially Halfdan Sørby, Lisa Balteskard, Per Pfeiffer, Kjell Magne Tveit, Reino Heikkilä, Nina Keldsen, Maria Albertsson, Hans Starkhammar and Hans Garmo.

All the former colleagues at the GI–section, Tone, Morten, Maria, Katarina, Masoud, Henrik, Mia, Mats, Tobias, Gisela, Agneta, Susanne, Anna, Tove, Sofie, Mona, Babro, Ing-Marie, Helena, Monica for all your good mood and spirit and professionalism, making the workdays enjoyable in spite of the difficult task.

Signe, Gabriella, Peter, Pia, Magnus, Eva, Lotta, Tina, Kjell, Annika, Christer, Sandra, Roger, Erika, Enrique, Anders, Per, Khairul and all the rest of my former colleagues for being such a bunch of warm hearted and extraordinary people.

Ulrik Ringborg, and later Sten Nilsson, head of the oncology clinic, for giving me the opportunity to do my research.

My friends Dan, Peter and Bim, for being who you are, you started medical school with me, we have all formed families, raised wonderful kids and grown older together, I hope it will continue so.

My parents Gunnar and Kerstin, for all support and help throughout life.

My wife Camilla, for invaluable support and love, and for your patience with me the last months.

And finally, our daughters Cornelia and Filippa, for being there and making everything worthwhile.
REFERENCES


42. C. Bokemeyer IB, J. T. Hartmann, F. G. De Braud, Volovat, J. Nippgen, C. Stroh, I. Celik, P. Koralewski. KRAS status and efficacy of first-line treatment of patients with

43. Van Cutsem E, Li, D’haens G et al. KRAS status and efficacy in the first-line treatment of patients with metastatic colorectal cancer (mCRC) treated with FOLFIRI with or without cetuximab: The CRYSTAL experience. J Clin Oncol 2008;26(May 20 suppl; abstr 2).

44. Maughan. "COIN: A Phase III trial comparing either CONTinuous chemotherapy plus cetuximab or INtermittent chemotherapy with standard continuous palliative combination chemotherapy with oxaliplatin and a fluoropyrimidine in first line treatment of metastatic colorectal cancer". ECCO/ESMO Congress. 2009:Abstract No:6LBA.


96. Hurwitz H. "Analysis of outcomes of patients with metastatic colorectal cancer (mCRC) treated with IFL with or without bevacizumab (BV) in a phase III clinical trial based on baseline risk". Proc Am Soc Clin Oncol. 2006:abstr 249.


