Radiation treatment in prostate cancer: balancing between tumor control and toxicity

Heemsbergen, W.D.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 8

Discussion
Chapter 8

1 Relationships between dose and toxicity

Dose effect relationships
A substantial part of this thesis concerns acute and late side effects, within the context of external beam radiotherapy of prostate cancer. The main focus is on gastrointestinal (GI) toxicity. In Chapter 4 the toxicity data are compared for the two dose arms of the Dutch randomized trial. In Chapter 2 and 3, toxicity is studied as a consequence of the specific dose received by organs at risk, especially the rectum and anal canal. In Chapter 5, late toxicity is studied as a result of dose volume effect relationships and acute toxicity.

Linking clinical data concerning side effects to individual risk factors and dose volume data, is an effective tool to learn more about the underlying dose effect relationships. Results of these analyses can be translated into guidelines for treatment planning in clinical practice, in order to minimize risks for toxicity as much as possible. In this light, it is obviously of great importance that the relevant complaints (i.e. which were indeed a result of the irradiation), are linked to the correct dose volume data (i.e. derived from the organ that is causing these complaints when it is disabled by irradiation). Furthermore, these dose effect relationships can be described separately for subgroups of patients, who are at a higher risk for developing side effects. One should however be careful with recommended dose constraints from other studies and institutions. It has been demonstrated that using different guidelines for the delineation of the (ano)rectum influences the dose volume histogram of the organ at risk, considerably. [1] This implies that dose constraints for relative volumes, achieved from studies with a different contouring of the organ at risk, can under- or over-estimate the risks. Another issue is the lack of comparable (toxicity) results in literature, as described in more detail in the paragraph on pitfalls in the comparison of results between studies (page 146).

Dose distributions in the anorectal wall
The Chapters 2 and 3 concern two studies in which we investigated the relationships between dose distributions and toxicity outcomes. For this purpose, we had access to the clinical and dosimetric data of a subset of the Rotterdam trial. In this trial, toxicity outcome for patients treated with rectangular fields were compared with conformal fields, in a patient population of 266 patients. [2] From our first study (Chapter 2), we learned that in order to investigate the relationship between the dose on the rectal wall and the endpoint late rectal bleeding, it was best to calculate relative
dose distributions (instead of absolute), to take dose to the wall only (instead of dose to the whole rectum including filling) and to investigate especially the dose to the rectum excluding the last 3 cm of the drawn rectal structure where the anal canal is located. Furthermore, the data suggest that the risk for bleeding keeps constantly rising with increasing dose rather than there is a kind of “threshold” value dividing the dose range into “low risk” and “high risk”. In Chapter 3, we investigated (in the same dataset) the relationship between the (relative) dose distributions in several parts of the rectal wall (including anal region), and a number of patient-reported symptoms (bleeding, mucus loss, rectal incontinence, increased stool frequency), in order to hypothesize which part of the rectal wall was most optimal related to each complaint. From this study we concluded that only bleeding correlated with dose, when the total structure (including anal region) was analyzed, whereas also other symptoms correlated with dose when only a superior or inferior part of the delineated organ at risk was evaluated.

The majority of studies concerning the relationship between irradiation of the rectum plus anal region and adverse events, focus on the relationship between the delineated rectum and the recorded EORTC/RTOG toxicity scores [3] Grade ≥ 2, with a main focus on rectal bleeding. Published data, in which the anal region and/or (part of the) rectum is analyzed in combination with other endpoints than bleeding, are limited. Koper et al. [2] reported also on this subject and divided dose volume information and toxicity into “anal” and “rectal” information. They could, however, not establish a significant relationship between dose parameters and the assigned toxicity scores. Al-abany and colleagues reported on significant dose volume effect relationships for the anal sphincter region in relation to fecal leakage. [4] Therefore they recommended limited volumes receiving intermediate doses levels (of 35 - 40 Gy) in the region of the anal sphincter.

**Consequential damage**

In Chapter 5 we investigated the relationship between acute and late gastrointestinal (GI) toxicity, including acute complaints and relevant dose parameters (concerning rectum, anal canal or both) in a multivariate analysis, for different endpoints (grade ≥ 2 toxicity: severe rectal bleeding, intermittent rectal bleeding, use of incontinence pads for fecal loss, use of steroids, stools ≥ 6 / day). We concluded that acute toxicity (i.e. acute proctitis) is an important independent predictor for late toxicity, with regard to increased stool frequency, use of incontinence pads, and intermittent bleeding. The dose parameters included in this analysis were derived from another published study of our group, in which we identified dosimetric parameters (anorectal, rectal and anal wall dose distributions) that correlated
with different late gastrointestinal complications. [5] The observation that acute GI toxicity is an independent predictor for late GI toxicity, has also been reported in a few other studies where dose volume information was also taken into account in a multivariate analysis. Wang and colleagues reported on acute diarrhea being an independent predictor for late proctitis, in patients treated with doses of 40 – 60 Gy for cervical cancer [6]. In a population of 743 prostate cancer patients, treated with 68 – 81 Gy, Zelefsky and colleagues found that acute GI toxicity Grade ≥ 2 was an independent predictor for late GI toxicity Grade ≥ 2. [7] Similar conclusions were drawn by Vargas and colleagues [8], in a study population of 331 patients. Zelefsky and colleagues also reported consequential effects for late GU toxicity.

For the prevention of late toxicity, by means of an optimal treatment plan, dose constraints are often aiming at high dose areas in the rectal wall, in order to limit the risk of late rectal bleeding. Taking into account that acute GI toxicity is an independent predictor for late GI toxicity, it would be meaningful to limit the risks for acute toxicity during treatment planning as well. As described by Peeters et al. [5], dose constraints to limit acute GI toxicity should not only aim at high doses, but also at intermediate doses. One step further to limit late toxicity by limiting acute toxicity would be to adapt the treatment itself. For instance, hypo-fractionation could attribute to less acute toxicity, and therefore less late toxicity. In 2007 Martin and colleagues [9] reported on acute toxicity rates for 92 patients treated with 60 Gy in 20 fractions of 3 Gy with a total treatment time of 4 weeks. Assuming an α/β of 1.5, the equivalent dose in fractions of 2 Gy would be about 77 Gy (on the prostate): the same dose as in the high dose arm of the Dutch trial. Maximum acute toxicity Grade 2 or higher was only 12 % for GI complaints and 25 % for GU complaints (RTOG scores). In the Dutch trial, these numbers were about 50 % for GI and GU complaints, in both arms (Chapter 4). Martin and colleagues also reported briefly on late toxicity according RTOG scales: at a median follow-up of 38 months actuarial late GI toxicity Grade ≥ 2 was 6 %, GU toxicity was 10 %. In the Dutch trial the cumulative Grade ≥ 2 GI toxicity at 3 years of follow-up, was 23 % and 27 % for the standard and high dose arm, respectively. For GU toxicity these numbers were approximately 30 % for both arms (Chapter 4).

The way of scoring toxicity is likely to be different between the study of Martin et al and the Dutch trial, despite the fact that the same scoring system was used; therefore a comparison should be made with caution. An important difference was that in the Dutch trial the patient questionnaires were used for scoring toxicity. In contrast, in the study of Martin et al., the physician assessed the toxicity. Nevertheless, the differences in reported toxicities are large and strongly suggest less bother for the patient in case of
Discussion

hypo-fractionation. It remains however the question if tumor control is indeed at the same level when comparing hypo-fractionated treatment with conventional fractionation. In the near future, more results on this subject will become available from randomized trials.

Gastrointestinal toxicity from a patient point of view

In Chapter 4 we reported on the (lack of) differences between the two dose arms with regard to Grade ≥ 2 toxicities; moderate differences in acute and late Grade ≥ 2 toxicities were found, with only a significant result for late rectal bleeding needing treatment and for late nocturia. Therefore we concluded that the dose escalation of 10 Gy is safe, because there is no relevant rise in moderate to severe toxicity. Another important question is to what extent a higher dose is acceptable for the patient, when it comes with more low grade toxicities or chronic complaints which significantly affect the quality of life of the patient in the remaining years of his life.

In the Dutch randomized trial, a large part of the patients reported their complaints on patient questionnaires, filled out during treatment and at each follow-up visit. Apart from the complaints, the patient was also asked if they were bothered by their complaints, grading their bother (for GI and GU complaint separately) as none, mild, moderate or severe. The available data on this bother score for GI complaints are summarized in Fig. 1. This Figure shows how many patients were “bother free” with regard to stools before radiation treatment and during a period of 4 years after treatment. The proportion of patients without bother is shown for the 68 Gy arm and the 78 Gy arm separately (for the latter group, patients planned with an advanced technique (IMRT) are excluded). The number of available questionnaires was 321 (68 Gy) and 289 (78 Gy) at year 1, down to 130 and 136 respectively, at year 4.

From Fig. 1 it is clear that after radiation treatment more patients are bothered with their stools: before RT about 90 % is free from complaints, after RT this percentage drops and increases again after 2-3 years. The fraction of patients indicating moderate to severe bother was at 3 years of follow-up: 9 % in the 68 Gy arm against 12 % in the 78 Gy arm. The difference between 78 – conventional technique and 78 Gy – advanced technique are significant at 1 and 2 year of follow-up ($p = 0.04$ and $p = 0.004$, respectively). Data of the small group of IMRT patients (not shown) show much less complaints: only 3 % reports “bother” at 3 years of follow-up. These results are, however, not significantly different.
Another issue related to the described bother of the patient, is which GI symptoms in particular are mostly related to this bother of the patient. The trial data set on “bother” and scored symptoms is, however, limited, because only part of the patients reported “bother”. At the department of Radiation Oncology of the NKI / AvL there is a much larger data set available on this subject, containing about 500 patients treated with 68-70 Gy in fractions of 2 Gy, between 1995 and 2002 with a total of about 4000 records. These data show that during follow-up visits about 30 % indicates to be bothered during stools because of GI complaints. This number of about 30 % is stable throughout the years up to 8 years of follow-up. Linear regression shows that all registered complaints contribute significantly to the ‘bother’ score. The main items contributing the most to the “bother” score are: pain, cramps, urge and uncontrolled loss of feces. The patient is less bothered by an increased stool frequency, blood loss and mucosal loss.

In a study by Koper et al [2], it was also observed that patients were most bothered by compliance related symptoms like soiling, fecal loss and urgency. Al-Abany and colleagues, who studied the relationship between irradiation of the anal sphincter and GI toxicity, came to similar conclusions: they reported that fecal leakage was the strongest factor associated with GI distress. [4]

A previous report on the Dutch trial about dose effect relationships [5], revealed a strong correlation between the mean dose to the anal canal
and fecal incontinence Grade ≥ 2. The reported cumulative incidence at 3 years for this endpoint was 5 % in the 68 Gy arm versus 10 % in the 78 Gy arm (Log Rank $p = 0.03$) in this data set of 641 evaluated patients. Although this difference was not (yet) significant ($p < 0.01$) it gives rise to concern, given the previous statements that fecal incontinence is one of the patient most feared side effects. For the small patient group treated in the high dose arm with the advanced IMRT technique, the estimated cumulative incidence of grade ≥ 2 fecal incontinence at 3 years was 5 % (data on the IMRT technique are planned to be published in the year 2008).

**Quality of Life scoring in the Dutch randomized trial**

In the Dutch randomized trial, the main focus concerning side effects was on GI and GU toxicity. A large part of the patients (n = 268) of the Dutch trial also filled out questionnaires on health related quality of life (HRQL) and a questionnaire on sexual items. The questionnaire on sexual functioning was analyzed in a separate analysis. Results were published in 2007 by van der Wielen *et al.* [10] They found a significant increase in the prevalence of erectile dysfunction (ED) after radiotherapy. In this study patients were censored at the moment they started with hormonal therapy, in order to estimate the effect of radiotherapy only. For the patients with no pre-RT complaints (72 % of the population), the cumulative incidence of ED was 38 % at 3 years. No significant differences were observed between the two dose arms.

Several published HRQL studies concluded that there were no differences for prostate cancer survivors in HRQL between different treatment strategies, including “watchful waiting”. Recently, a Dutch survey study with 780 respondents, reported worse HRQL for patients treated with hormonal therapy scored on the 36-item Short Form Health Survey (SF-36) compared to other treatment options (radiotherapy, watchful waiting, prostatectomy). Comparing scores on the subscales (e.g. bodily pain, social functioning) of the radiotherapy group with an age-matched population, no clinical meaningful differences were found. [11] Despite the lack of differences in HRQL scores in prostate cancer survivors compared to control groups, they do, however, report significantly more problems with sexual functioning, stools and urinating. Apparently, these problems are not translated into worse HRQL scores.
Chapter 8

2 A higher treatment dose in prostate cancer

Outcome of the Dutch randomized trial
In Chapter 6 the main results of the Dutch randomized trial (68 Gy vs. 78 Gy) are presented. In a study population of 669 patients with a median follow-up of 51 months, we found a significant improvement in freedom from (biochemical and clinical) failure (FFF) for the high dose arm, according the ASTRO definition (3 consecutive rising PSA’s, backdated to midpoint between first rise and last non-rising value). The overall reported adjusted Hazard Ratio (HR) in our study, adjusted for treatment group, age, institution and hormonal therapy was 0.74 (\( p = 0.02 \)) for the total trial population in the intention-to-treat analysis. We also evaluated the dose effect for the actually given dose (below 73 Gy versus above 73 Gy) which indicated a substantial larger benefit for the high dose group with a mean treatment dose of 77.9 Gy (versus 67.9 Gy): an adjusted HR of 0.66 (\( p = 0.001 \)). About 10 % of the high dose arm was treated at a lower dose, mainly because dose constraints could not be met. For the endpoint clinical relapse (loco regional recurrence, metastasis or initiation of hormonal treatment), no differences were found between the two arms.

As discussed in Chapter 6, the conclusions of our trial confirm the conclusions drawn in other previous studies on the same subject. We compared our results in particular with the results of the M.D. Anderson trial which had a similar study design, comparing 70 Gy with 78 Gy in a study population of 301 patients. [12,13] The main results (freedom from failure according the ASTRO definition) were published in 2002. They found an adjusted HR of 0.59 (\( p = 0.02 \)) for the total group and they reported that the dose effect could only be attributed to the subgroup with an initial PSA (iPSA) above 10 ug/L (no dose effect for iPSA \( \leq 10 \) ug/L). Because our trial was not designed for subgroup analysis (in fact, neither was their trial) we did not draw definitive conclusions for these subgroups. However, because several authors have reported a lack of dose effect for patients with an iPSA \( \leq 10 \) ug/L, an explorative analysis on this subject was performed on the data of the Dutch randomized trial. The results are described in the paragraph on dose effects for subgroups based on initial PSA (page 148).

Recently, the results of another large randomized clinical trial with a similar study design were published by Dearnaley et al: the MRC RT01 randomized trial. [14] Also in this trial the main conclusion is that a higher dose (in this study 74 Gy versus 64 Gy) results in better freedom from clinical and biochemical failure. A more detailed comparison of the results of
the Dutch trial and the MRC trial is described in the paragraph “Comparison of the Dutch trial with the MRC RT01 trial” (page 147).

**The Phoenix definition for freedom from biochemical failure**

In the Dutch trial the applied definition of biochemical failure (as an early surrogate endpoint for clinical failure) was the recommended guideline of the ASTRO at that time: three consecutive rises in PSA level with the date of failure between the first rise and the last non-rising value. [15] In 2006 a report was published concerning the evaluation and revision of these ASTRO recommendations. The investigators at this meeting concluded that an alternative definition was more appropriate, especially when follow-up is limited and hormonal therapy is prescribed: a rise of 2 μg/L above the post-RT nadir (the Phoenix definition). [16] They also stated that “3 consecutive rises backdated to the date between first rise and nadir”, was not an appropriate definition when hormonal therapy is prescribed, especially when (additional) HT and no HT are compared, and when the frequency of follow-up visits is low. They also recommended that freedom from failure rates should be reported at 2 years short of the median follow-up, in order to avoid the problems associated with backdating. Because biochemical failures are backdated according the definition of 3 rises, Kaplan–Meier curves tend to shift each time more follow-up becomes available, which makes it tricky to compare rates between different studies.

The outcome of the Dutch trial has been recalculated using the Phoenix definition, as briefly mentioned in the discussion of Chapter 6. It shows better tumor control in the 78 Gy arm which is not significant ($p = 0.3$, Log Rank test). Comparing the two dose arms according actual dose (< 73 Gy, mean of 66.9 Gy versus > 73 Gy, mean of 77.9 Gy), this $p$ value is 0.13. In Fig. 2 the (smoothed) Kaplan–Meier curves are shown for freedom from (biochemical or clinical) failure according the ASTRO definition and for the Phoenix definition. The dose arms defined according the true dose contain 358 (<67.9 Gy subgroup) and 306 (77.9 Gy subgroup) patients. In order to facilitate a fair comparison, the ASTRO curves are shown without backdating. When comparing the definitions, the moment in time that the curves start to separate is 6 months apart: at about 2.5 years for the ASTRO definition and at about 3 years for the Phoenix definition. Furthermore, the dose arms are separating a bit faster according the ASTRO definition and the levels of tumor control drop faster after 4 years, especially in the 67.9 Gy subgroup. From the latest update of the Dutch trial [17], with more follow-up available, it is known that there is also a significant difference in FFF between the two randomization arms according the Phoenix definition (results not shown).
Pitfalls in the comparison of results between studies

It would be interesting, from a research point of view as well as from a clinical point of view, to compare the levels of tumor control and toxicity of the Dutch trial with the results found in other published studies with similar irradiation schemes. Also a comparison with historical data, when for instance elective irradiation of the pelvic region was applied and lower doses were prescribed, would be interesting. It appears however to be hard to make valid comparisons.

First of all, the endpoint ‘tumor control’ is defined in a different way throughout the years. The ASTRO committee has made an effort to tackle this problem by the publication of guidelines for the definition of “freedom from clinical and biochemical failures” for randomized trials. Another issue is that study populations are often not comparable with regard to the prognostic features (e.g. T stage, Gleason score, initial PSA level). Moreover, throughout the years other characteristics of the tumor are presented (e.g. historical data do not report on Gleason scores and PSA). A third problem is that described subgroups are often based on different criteria, which applies to older studies as well as to recent published data. And even when study populations are comparable and subgroups are defined according similar criteria, there remain other factors that impair a fair comparison of the outcome: a different frequency of follow-up visits, a different policy (e.g. guidelines for imaging, biopsy) and different criteria for

Fig. 2. Freedom from failure for the dose arms based on actual dose, according to the ASTRO definition (without backdating) and according the Phoenix definition.
the initiation of hormonal therapy after RT (which is in most studies defined as “failure”) can all influence the estimated failure rates. In When the ASTRO definition of three consecutive rises is used as an endpoint, also a different median follow-up period between studies will affect the comparisons of estimated freedom from failure, because ‘backdating” is applied, affecting these estimates. Also in the Dutch trial such features are present: a different level of freedom from failure was noticed between the hospitals, and this was for a large part due to the fact that one hospital could not meet the protocol guidelines for the frequency of follow-up visits.

With regard to comparing toxicity levels, similar problems arise. Many studies apply RTOG / EORTC toxicity scales which would make it in principle easy to make straight-forward-comparisons. However, the interpretation of these toxicity scales can differ considerably. Also differences in the frequency of follow-up visits, and the policy in evaluating the situation, have an effect on toxicity scores. For instance, when patients fill out questionnaires, much more complaints are registered then when the information on complaints is only noted by the physician. Recently, Garg and colleagues reviewed a large amount of data on radiation proctopathy in prostate cancer treatment. [18] They encountered many problems during the comparison of the data from different studies.

In conclusion, comparisons of results on toxicity and tumor control should be made very carefully. No need to argue that this is an important drawback in the general search for the “optimal radiation treatment” for prostate cancer.

Comparison of the Dutch trial with the MRC RT01 trial

For the Dutch trial the overall reported adjusted Hazard Ratio (HR) was 0.74 for freedom from failure in the 78 Gy arm compared to the 68 Gy arm\( (p = 0.02) \) in the intention-to-treat analysis (median follow-up of 51 months). Dearnaley and colleagues reported an overall HR of 0.67 for the high dose group in the MRC RT01 trial (74 Gy vs. 64 Gy) after a median follow-up of 63 months, which is in general agreement with our results. There are however some relevant differences between the studies. Apart from the difference in the described dose, differences in the study design were: in the Dutch trial patients with stage T3b were included (14 % of the population) whereas only stage ‘T1b-T3a’ patients were allowed in the MRC trial. In the Dutch trial only part of the patients received hormonal therapy (HT) prior to RT (21 %, mainly moderate to high risk patients) whereas all patients received HT in the MRC trial. Another issue is that the high dose was prescribed to the total PTV (including seminal vesicles (SV)) for the MRC trial whereas in the Dutch trial the dose to the SV was 0 Gy
(patients with favorable tumor characteristics), 50 Gy or 68 Gy and not depending on the randomization arm (except for a small group of high risk patients who received 78 Gy to the SV in the high dose arm). Another relevant difference between the two studies was a different definition of biochemical failure. In the MRC RT01 study the definition was: a PSA concentration greater than 2 ng/L and at least a 50 % increase compared to the nadir. This definition is close to the “Phoenix definition”, as described earlier (page 145). When the Phoenix definition is applied to the Dutch trial data, the estimated HR for FFF becomes 0.86 (intention-to-treat) and 0.81 according the actual given dose (< 73 Gy).

Both trials published results for risk subgroups (low, intermediate, high risk) according Chism et al. [19] Comparing the results for these subgroups, the estimated hazard ratios (HR) are (Dutch trial and MCR trial, respectively): 1.2 vs. 0.8 (low risk), 0.5 vs. 0.7 (intermediate risk) and 0.8 vs. 0.6 (high risk). When we look at the HR according the actually given dose for the Dutch trial data (< 73 Gy >) the HR becomes 0.8, 0.5 and 0.7, respectively. The latter shows that the HR estimate for the “low risk” subgroup does indicate a (not significant) dose effect, in contrary to the results according the intention-to-treat analysis. A significant interaction between dose and risk group is however not present (i.e. the dose effect is not significantly different within the risk groups) because both trials were not designed for reliable subgroup analyses. For this and other reasons, a meta-analysis of several randomized trials on this subject would be very useful.

**Dose effect for subgroups based on initial PSA**

In the Dutch randomized trial, 39 % of the patients had an initial PSA level (iPSA) of 10 ng/L or below. It has been observed in a number of studies that patients with an iPSA ≤ 10 ng/L do not profit from higher treatment doses. Studies concerning patient cohorts as well as data from a randomized trial, suggest this. Pinover and colleagues [20] published in the year 2000 that patients with a PSA ≤ 10 ng/L do not profit from dose escalation, except for the “poor prognosis” subset (at least T2b and/or Gleason at least 7 and/or perineural invasion). Their dataset consisted of historical cohort of about 500 patients with an initial pretreatment PSA of 10 or below. In 2003, the results of the M.D. Anderson trial were published [12]: they reported no dose response for the subgroup with a pretreatment PSA ≤ 10 ng/L. In 2007, an update of the trial results was presented in which they confirmed these observations. [13]

A first analysis of our data shows no dose effect for patients with an iPSA of ≤ 10 ng/L; a HR of 0.97 for the endpoint FFF. When we analyze the data in detail, evaluating the actually given dose and evaluating more iPSA
subgroups, we come however to different conclusions. When we look at the actual given dose (66.9 Gy group and 77.9 Gy group), the adjusted HR becomes 0.73 ($p = 0.2$) in the PSA $\leq 10$ µg/L subgroup. Furthermore, when we analyze above and below the median in this subgroup (iPSA = 6.8 µg/L) a dose effect is clearly suggested (actual dose) for the patients with a PSA > median (HR = 0.63, $p = 0.2$) whereas this HR is 0.83 for the subgroup below the median iPSA.

![Graph](image)

**Fig. 3.** Freedom from failure according the ASTRO definition for the “actual dose” arms of the trial. Data are shown for the total patient group ($n = 664$) and for the patient group with an iPSA between 8 and 18 µg/L ($n = 294$).

A further explorative analysis of our data, show the most optimal dose effect (based on actual dose) for patients with an iPSA between 8 and 18 µg/L (294 patients, 44 % of the total group), as illustrated in **Fig. 3** (adjusted HR of 0.49, $p = 0.001$). Above and below these cutoffs, the HR is in the range of 0.8-0.9.
In conclusion, our data do not fully support the findings in other studies that patients with an iPSA of $\leq 10\, \mu g/L$ do not profit from dose escalation. For very low iPSA’s this seems true but for iPSA levels above a certain level (probably between 6-8 $\mu g/L$) the data suggest that dose escalation is useful. The trial data suggest a dose response for patients with an iPSA roughly within the range of 6-20 $\mu g/L$ (and an optimal dose response in the range of 8-18 $\mu g/L$). Below and above these cutoff levels, only a poor dose response was established in the data set.

Of note is first, the explorative nature of these analyses. For instance, the exact cutoffs for “a low iPSA” and a “rather high PSA”, as well as for the most favorable group, cannot be established accurately from our data alone. Furthermore, different dose effects in subgroups, based on a combination of iPSA and other (un)favorable characteristics, as suggested by Hanks et al. [21], cannot be excluded or confirmed in our dataset, because the statistical power is too limited. As suggested before, a meta-analysis of several trials would be useful. A final comment is that the determination of the iPSA levels can differ between different assays, and is probably nowadays more accurate then 10-20 years ago.

**Suboptimal treatment and tumor control**

In Chapter 7 we demonstrated that our treatment protocol is not optimal for a subgroup of patients, with a relatively large rectum on the CT scan and who developed diarrhea during treatment. Data show a decrease in tumor control for these patients, when they have a relatively large risk for involvement of the seminal vesicles and for extra capsular extension of the tumor. Also the prescribed margin of 0 mm for the boost of 10 Gy at the dorsal (rectal) site, was probably not optimal. When comparing the data (Kaplan Meier curves), the gain in tumor control by adding 10 Gy is negligible for the patients with the full rectum and diarrhea: the estimated levels of FFF in the 78 Gy group for these patients is approximately the same as for patients treated at 68 Gy. In Chapter 7 we suggested that the problem of a prostate position on the CT scan not being representative for the position during treatment, can be solved by improving the clinical protocol (e.g. laxatives prior to the CT scan and/or daily imaging of the prostate). From further research on this subject we learned, however, that the underlying mechanisms leading to the suboptimal treatment are almost certain more complicated and cannot be solved completely by improved image guidance during treatment.

Based on the results described in Chapter 7, we have tried to simulate the observed drop in tumor control, using the data in our volumetric trial database containing the CT scans, delineations and dose distributions. For this purpose we used a TCP (tumor control probability) model to fit our FFF
data. It appeared that if we assume that the drawn CTV is correct, it is not possible that the tumor control drops so dramatically due to a modest shift caused by a relatively empty rectum during RT for the patients at risk (i.e. with a full rectum on the CT and diarrhea during treatment).

Our hypothesis based on these findings, was that probably the delineation of the prostate on the CT scan is not the optimal clinical target volume for a number of patients, which could be due to extra capsular invasion (ECI) which cannot be seen on the CT scan. Further explorative analysis of our data supported the hypothesis that the CTV was probably too tight for a number of patients: the dose to the rectum (especially high dose areas) appeared to be a very significant predictor for outcome, which is strongly suggestive for the presence of tumor cells in the area around the rectal wall which is also the area at risk for ECI. An example of visible ECI on an MR scan is shown in Fig. 4: a significant amount of extra capsular tumor growth is identified on the MR scan and as a consequence the GTV is drawn outside the prostate (this scan concerned a patient of 75 years old with a clinical T3a tumor, Gleason 8 and an iPSA of 9 μg/L).

An optimal CTV possibly also includes a number of lymph nodes for a part of the patient population. Radiotherapy of the regional lymph nodes, for patients with an estimated risk for node involvement of > 15 %, has been proposed as a superior treatment, although other research groups did not find improvement of tumor control when pelvic nodes were included in the CTV. [23,24] An estimated node involvement of > 15 %, according the Partin tables [25], applies to about half of the patient population in the Dutch trial (roughly: T2 tumors with PSA > 10 μg/L and Gleason > 6 and all T3/T4 tumors). To include lymph nodes in the CTV without irradiating the whole pelvis is complicated [23], because the anatomic position of the lymph nodes varies largely between patients and is difficult to identify on CT images. A whole pelvic irradiation solves this problem, but this is of course a major limiting factor for the possibilities of dose escalation.
3 Further optimization of radiation treatment

The benefit of a higher treatment dose
So far, there is no evidence that a radiation dose above 70 Gy improves overall survival in prostate cancer patients, compared to doses in the range of 64 Gy - 70 Gy. An escalated prescribed dose, in the range of 74-80 Gy, is superior to a prescription dose in the range of 64-70 Gy, for the average prostate patient with regard to tumor control in terms of biochemical or clinical relapse (FFF). There are also indications that the risk for clinical relapses (FFCF) decreases with a higher treatment dose. In the recent update of the M.D. Anderson trial [13], a superior FFCF was reported for the 78 Gy arm, compared to 70 Gy (93 % versus 85 % at 8 years of follow-up, \( p = 0.01 \)). In the MRC RT01 trial [14] a Hazard Ratio of 0.69 (\( p = 0.06 \)) was reported, for the risk of clinical failures in the high dose arm (74 Gy versus 64 Gy). In the Dutch trial no such indications were found. Whether a higher dose also leads to less prostate cancer-related death, remains unknown until more data on this subject become available (more follow-up from the randomized trials). It is likely that only a subgroup of patients with localized prostate cancer benefits from a higher prescription dose in terms of improved survival. Definitive conclusions regarding subgroups are probably possible in the future when meta-analyses of the randomized trials are performed.
**Discussion**

**The benefit of a correct determination of the CTV**

The results of our analyses on geometric misses (Chapter 7) stress the importance of a careful and correct determination of the location of the clinical target volume (CTV), during treatment planning and the treatment itself. As described in Chapter 7, the determined CTV was probably too tight for a number of trial patients, due to the presence of extra-capsular invasion (ECI). The Partin tables report high risks of ECI: in general, only tumors with a low iPSA (< 2.5) or a low Gleason score (< 6) have a risk below 30 %. Otherwise, estimated risks are up to 60 % for T1c-T2 tumors. Therefore, it would make sense to translate the risks for ECI into adequate coverage during treatment planning. This was also recommended in a recent publication of Boehmer et al. [24], on behalf of the EORTC radiation Oncology Group. Another issue with regard to the determination of the CTV, is the inclusion of lymph nodes into the treatment volume. At this moment, there is, however, no consensus for which patient groups this would improve outcome. A general conclusion is, that more knowledge about sub-clinical disease is needed, to improve treatment in “localized” prostate cancer.

**The benefit of less toxicity**

From the Dutch randomized trial, and from other studies, there are strong indications that advanced treatment planning (limiting the dose to organs at risk), keeps acute and late toxicity at lower levels. Because the experienced toxicity after radiotherapy is considerable when conventional planning techniques are applied, further optimization of treatment planning is beneficial for a large patient group. Tighter dose gradients are, however, also associated with increased risks of marginal misses. Therefore, a reliable CTV determination, and a translation of all relevant errors into the correct CTV-PTV margins, is becoming more and more an issue in advanced treatment planning.

**Other research issues**

Apart from the issues discussed in this thesis, there are other urgent research questions with respect to an optimal treatment of prostate cancer. Main issues discussed in literature are for instance: 1) what are the optimal dose and fractionation schemes, and 2) what is the optimal hormonal therapy in combination with RT. Therefore, the external beam treatment of the (near) future is possibly a completely different treatment than it is nowadays, leading to a renewed need for research on toxicity, dose effect relationships and tumor control.
4 Balancing between tumor control & toxicity: Conclusions

We have come to the following conclusions with regard to treatment of localized prostate cancer:

1) With the current standards in radiotherapy planning, a dose up to 78 Gy in daily fractions of 2 Gy can be delivered to the patient without unacceptable risks of severe toxicity.

2) An escalated prescribed dose in the range of 74-80 Gy is superior to conventional doses of 64-70 Gy for the average prostate patient with regard to tumor control in terms of biochemical or clinical relapse. Whether this is also associated with less prostate cancer related death remains unknown until more data on this subject (more follow-up from the randomized trials) becomes available.

3) There are strong indications that not every patient will profit from a higher treatment dose. A major step in understanding when a higher dose is favorable and when it is not could be achieved by a meta-analysis of the trials on this subject.

4) There is a need for a more adequate determination of the CTV, i.e. the GTV and the (sub-clinical) spread of tumor cells in the loco regional tumor area. This calls for a more optimal integration of knowledge of different disciplines (i.e. radiotherapy, pathology, radiology). An MR scan for patients at risk for large ECI could be useful in clinical practice.

5) With the current knowledge, we can identify on beforehand patients at a higher risk for RT-related toxicity. There is a need for a broad discussion in the literature how to incorporate this in the treatment decision and planning.

6) Apart from toxicity risks for severe rectal bleeding, also the risks for side effects affecting the patients quality of life in the remaining years (like fecal incontinence), should be included in the planning and decision making process. This implies evaluation of the dose to the anal region.

7) Dose constraints during treatment planning should also aim at preventing acute toxicity, because acute toxicity is an important independent risk factor for late toxicity.
References


