Evaluation of uncertainty predictions and dose output for model-based
dose calculations for megavoltage photon beams

Jörgen Olofsson and Tufve Nyholm
Department of Radiation Sciences, Radiation Physics, Umeå University, SE-901 87 Umeå, Sweden

Dietmar Georg
Div. Medical Radiation Physics, Department of Radiotherapy and Radiobiology,
Medical University Vienna-AKH, Währinger Gürtel 18-20, A-1090 Vienna, Austria

Anders Ahnesjö
Department of Radiation Sciences, Radiation Physics, Umeå University, SE-901 87 Umeå, Sweden
and Nucletron Scandinavia AB, P.O. Box 1704, SE-751 47 Uppsala, Sweden
and Department of Oncology, Radiology and Clinical Immunology, Uppsala University, Sweden

Mikael Karlsson
Department of Radiation Sciences, Radiation Physics, Umeå University, SE-901 87 Umeå, Sweden

(Received 20 December 2005; revised 23 February 2006; accepted for publication 27 April 2006; published 26 June 2006)

In many radiotherapy clinics an independent verification of the number of monitor units (MU) used
to deliver the prescribed dose to the target volume is performed prior to the treatment start. Traditionally this has been done by using methods mainly based on empirical factors which, at least to
some extent, try to separate the influence from input parameters such as field size, depth, distance,
etc. The growing complexity of modern treatment techniques does however make this approach
increasingly difficult, both in terms of practical application and in terms of the reliability of the
results. In the present work the performance of a model-based approach, describing the influence
from different input parameters through actual modeling of the physical effects, has been investiga-
ted in detail. The investigated model is based on two components related to megavoltage photon
beams; one describing the exiting energy fluence per delivered MU, and a second component
describing the dose deposition through a pencil kernel algorithm solely based on a measured beam
quality index. Together with the output calculations, the basis of a method aiming to predict the
inherent calculation uncertainties in individual treatment setups has been developed. This has all
emerged from the intention of creating a clinical dose/MU verification tool that requires an absolute
minimum of commissioned input data. This evaluation was focused on irregular field shapes and
performed through comparison with output factors measured at 5, 10, and 20 cm depth in ten
multileaf collimated fields on four different linear accelerators with varying multileaf collimator
designs. The measurements were performed both in air and in water and the results of the two
components of the model were evaluated separately and combined. When compared with the
respective measurements the resulting deviations in the calculated output factors were in most
cases smaller than 1% and in all cases smaller than 1.7%. The distribution describing the calcula-
tion errors in the total dose output has a mean value of −0.04% and a standard deviation of 0.47%.
In the dose calculations a previously developed correction of the pencil kernel was applied that
managed to contract the error distribution considerably. A detailed analysis of the predicted uncer-
tainties versus the observed deviations suggests that the predictions indeed can be used as a basis
for creating action levels and tracking dose calculation errors in homogeneous media. © 2006
American Association of Physicists in Medicine. [DOI: 10.1118/1.2207316]

Key words: Output factors, uncertainties, pencil kernel, MLC, MU verification

I. INTRODUCTION

Treatments in external radiotherapy are gradually becoming
more complex and make use of an increasing number of
irregularly shaped radiation fields, as in intensity modulated
radiotherapy (IMRT). The single most important piece of
new equipment that has made this development possible is
the multileaf collimator (MLC). This was originally a tool
developed in order to replace individual shielding blocks, but
the fast and simple handling of differently shaped fields has
been an incentive to develop the present methods used to
plan and deliver IMRT.

The design solutions that currently are provided by major
manufacturers of linear accelerators, regarding the location
of the MLC in the treatment head, suggests that the evolution
still not has converged. Some manufacturers have chosen to
replace one of the collimating jaw pairs while others have
preferred to add the MLC as a third level of collimators. When implementing calculation models for beam output these differences in design are of great importance. The position of additional backup collimators in relation to the field edges is also an issue that can have a significant influence on contributions related to treatment head scatter and on penumbra shaping.

In most clinics some form of independent verification of the number of monitor units (MU) necessary to deliver the prescribed dose is performed. Together with the increasing number of treatment fields comes however a growing problem associated with the quality assurance (QA) procedures handling individual treatment plans. The MU verification is commonly performed through an independent calculation of dose per MU using factor-based models, originally designed for rectangular field shapes.\textsuperscript{1,2} But as the total dose is divided into many small portions, coming from irregularly shaped subfields (segments), factor-based models will require substantial amounts of work and possibly also introduce considerable uncertainties.\textsuperscript{3}

There is therefore a need for a more general tool for independent dose/MU verification. Such a tool should be quick and easy to use, both in terms of beam commissioning and in daily use dealing with individual treatment plans. Another vital demand is that this tool should be completely independent from dose calculation data coming from, or used as input to, the treatment planning system (TPS). However, the issue of easy handling does not reduce the demands on calculation accuracy. Errors in the order of a few percent should be detectable. Ideally each calculation result should also come with a corresponding uncertainty estimation that can facilitate appropriate action levels in the clinical QA routine.\textsuperscript{3}

Studies on the influence from irregular MLC fields on photon beam output have previously been presented in the literature.\textsuperscript{5–7} These investigations describe calculations that are built on physical modeling rather than strict factorization of measured quantities, which also means that there is a more or less inherent ability to deal with features like irregular field shapes. Some works have been focused on irregular MLC fields and output factors in air, i.e., head scatter factors, rather than the total output.\textsuperscript{8–11}

However, in most cases the investigations have only covered one specific type of multileaf collimator, which means that the general applicability of the algorithm has not been tested. Georg et al.\textsuperscript{12} performed a study similar to the present one (but only at 10 cm depth) to see whether more traditional and empirical methods, such as equivalent squares, could be used for MU verification also in highly irregular MLC fields. The conclusion was that for irregular fields where central parts of the flattening filter were blocked by leaves large discrepancies (up to 6%) exist, which is unacceptable in clinical situations.

In the present work we have tested and evaluated a model-based dose/MU verification tool following the guidelines and ideas presented above. The investigation covered three treatment depths in several megavoltage photon beams. The field shapes and sizes were selected in order to thoroughly test the ability of the calculation model to deal with any kind of open irregular photon fields, except for small beams with field edges located closer than 2 cm from the point of interest. The motivation for this restriction was to exclude penumbra related issues, such as lateral nonequilibrium of charged particles and focal source distributions. The beams were delivered using different linear accelerators representing the most common types of treatment units and MLCs used clinically today. The evaluated dose calculation model includes a method describing the calculation uncertainties associated with individual field setups and the ability to create valid uncertainty predictions was analyzed by comparing the predictions with the actual deviations found in the results.

II. THEORY

A. Energy fluence calculations

The experimental quantity related to the total energy fluence $\Psi$ in air for a given position $x$ $(x, y, z)$ and field aperture setting $A$ per dose monitor signal $M$ can, after normalization under reference conditions, be described as

$$OF_{\text{air}}(x; A) = \frac{\Psi(x; A)}{M(A)} = \frac{\Psi(x_{\text{ref}}; A_{\text{ref}})}{M(A_{\text{ref}})},$$

where the reference point $x_{\text{ref}}$ is equal to the isocenter point in this case. The output factor in air ($OF_{\text{air}}$) was calculated in agreement with the principles previously outlined by Olofsson et al.\textsuperscript{13} Hence, the photons leaving the treatment head are emitted by three different sources (when no beam modulator is used)

$$\Psi(x; A) = \frac{\Psi_d(x; A) + \Psi_e(x; A) + \Psi_c(x; A)}{M_d + M_e + M_c(A)}.$$

The indices $d$, $e$, and $c$ in Eq. (2) represent direct (focal), extra-focal, and collimator contributions, respectively. These three sources also influence the dose monitor signal, but in this case it is only the part that comes from the variable collimators downstream from the monitor chamber ($M_c$) that varies with the collimator setting ($A$).

In the present work a few modifications of the algorithm for calculation of $\Psi_c$ have been introduced (even though the resulting effects from this generally are at a subpercent level). The extra-focal source distribution has been changed from a pyramid to a Gaussian distribution\textsuperscript{11,14} in order to model the phenomenon more accurately. This provides a rotation symmetric representation, which is in better agreement with the circular radiation field that exits the primary collimator. Furthermore, the angular correction factor $c_{\text{ang}}(\phi)$ for Compton scattering effects in the extra-focal source (originally parametrized by Ahnesjö\textsuperscript{15}) was previously based on the central visible point of the source. In the present implementation the Gaussian extra-focal distribution is represented by a $203 \times 203$ matrix and the scattering angle from each one of those 41 209 matrix elements toward each of the energy fluence sampling points is used to take this effect into account. The finite matrix result of $203 \times 203$ elements
was set in order to always avoid dose calculation errors larger than 0.1% that are caused by the numerical source integration.

Finally, in the current version the description of all variable collimators is generalized to model thick collimators. This is done by using three discrete layers instead of just one layer located at the inner edge of the actual collimator. The three layers, where each one represents one third of the collimating properties that are associated with the respective collimator, are located at the inner and outer collimator edges plus one in the middle. This means that the geometrical thickness of the collimators is more adequately modeled, which offers a more accurate ray tracing of the photons as they travel down through the treatment head geometry.

Consequently, the extra-focal contribution $\Psi_e$ was calculated as

$$\Psi_e(x;A) = \sum_{k=1}^{203} \sum_{l=1}^{203} \left[ \prod_{u=1}^{U} \prod_{v=1}^{V} t_{u,v}(x,k,l;A) c_{\text{ang}}(\phi) e_{k,l} \right]$$

$$t_{u,v} = \begin{cases} 
1 & \text{when no intersect with collimator layer (u,v)} \\
\frac{1}{\eta_u} & \text{when intersect with collimator layer (u,v)}
\end{cases},$$

where the sums over $k$ and $l$ represent the elements in the Gaussian source matrix, $u$ counts over the collimators in the treatment head, and $v$ corresponds to the number of discrete layers used to describe each collimator (in this case 3). $t_{u,v}$ is the transmission associated with each collimator layer, which is 1 when the ray from source element $(k,l)$ toward $x$ does not intersect layer $(u,v)$ and one third of the collimator transmission $\eta_u$ when it does intersect, $c_{\text{ang}}$ is the angular correction factor based on scattering angle $\phi$ between element $(k,l)$ and the calculation point $x$, and $e_{k,l}$ is the amplitude of source element $(k,l)$.

Perturbation in the dose monitor signal due to collimator backscatter ($M_s$) as well as forward directed collimator scatter ($\Psi_s$) was modeled according to Olofsson et al. This was also the case for the scheme used when characterizing the different head scatter components individually for each of the included photon beams, which means that the input data consisted of measured OF$_{\text{air}}$ from ten square fields.

### B. Dose calculations

The basic principles for the pencil kernel algorithm used to calculate the total output OF$_{\text{total}}$, i.e., the normalized dose per given MU, are described by Nyholm et al. The parametrization of the pencil kernels originates from the work by Ahnesjö et al., but in the implementation used here it is solely based on the beam quality index TPR$_{20,10}$ which makes the beam commissioning extremely fast and simple. This kernel model has also been expanded by Nyholm et al. in order to include the off-axis softening effect present in most megavoltage photon beams. The lateral shift in beam quality is better described in terms of half value layer (HVL) than TPR$_{20,10}$ and therefore HVL is the basis for characterizing this effect. A correction for systematic errors in the kernel characterization, derived from a statistical analysis based on a large number of measurements in clinical photon beams, was also included in order to further improve the accuracy.

The conversion from OF$_{\text{air}}$ to total beam output OF$_{\text{total}}$ was performed through numerical integration of pencil kernels in concentric circles over a plane perpendicular to the beam direction at the calculation depth $d$. This summation includes all dose-contributing sectors around the calculation point, weighted by the corresponding energy fluence.

$$\text{OF}_{\text{total}}(x,d;A) = \frac{D(x,d;A)/M(A)}{D(x_{\text{ref}},d_{\text{ref}};A_{\text{ref}})/M(A_{\text{ref}})} = M(A_{\text{ref}}) \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} (\Psi_{i,j,A}(\phi) \int_{r_{\text{in}}(p/p)(r,d,HVL_{i,j})}^{r_{\text{out}}(p/p)(r,d,HVL_{i,j})} 2\pi r \, dr)_{x_{\text{ref}}}}{\sum_{i=1}^{I} \sum_{j=1}^{J} (\Psi_{i,j,A}(\phi) \int_{r_{\text{in}}(p/p)(r,d_{\text{ref}},HVL_{i,j})}^{r_{\text{out}}(p/p)(r,d_{\text{ref}},HVL_{i,j})} 2\pi r \, dr)_{x_{\text{ref}}}}. \tag{4}$$

where index $i$ represents the different radii and $j$ the angular sectors for each radius ($J=15$ was used for all radii), $p/p$ is the pencil dose deposition kernel characterized by the average half value layer (HVL$_{i,j}$) in the sector. The basic principles are illustrated graphically in Fig. 1.

Prior to the dose summation for OF$_{\text{total}}$ in Eq. (4) the energy fluence values $\Psi_{i,j,A}$ were calculated for each sector and radius $(i,j)$ using the same algorithm as for the primary calculation point itself. The sampling algorithm was implemented so that the selection of radii was affected by the importance given by the radial kernel properties and the field shape, as illustrated in Fig. 1, although the number of sampling points per radius was kept constant at 60 (i.e., four points per sector). Hence, extra radius intervals were automatically added in regions including considerable energy fluence gradients in order to eliminate calculation errors emanating from insufficient sampling around the calculation point.

Finally, the dose at $x$ and depth $d$ per normalized energy fluence in air from field setting $A$, i.e., the effect of dose deposition in the medium, is denoted $q$. Consequently $q$ was given by

$$q(x,d;A) = \frac{\text{OF}_{\text{total}}(x,d;A)}{\text{OF}_{\text{air}}(x;A)}. \tag{5}$$
All results that are explicitly presented in this study refer to the isocenter point, i.e.,  \( x = x_{\text{ref}} \).

### C. Uncertainty estimations

In the uncertainty estimations dealing with calculated \( \text{OF}_{\text{air}} \), the method suggested by Olofsson et al.\(^{13}\) has been used. For each of the photon beams under study the basis (denoted \( s_{\text{air},0} \)) for this estimation is the standard deviation of the errors associated with the square fields that were used to individually characterize the model, or a minimum value of 0.15% whichever is larger. Additional uncertainty is then added by square summation, for different treatment situations. The size of the uncertainty associated with irregular MLC shaped fields (denoted \( s_{\text{air,irreg}} \)) was not known prior to this investigation and therefore a value was derived empirically by means of the previously described short field uncertainty \( s_{\text{air,short}} \). This procedure was performed through a square subtraction of \( s_{\text{air,short}} \) from the measured standard deviation in air (\( \sigma_{\text{air}} \)), averaged over the ten investigated photon beams, where the results were considered as being \( s_{\text{air,irreg}} \). In order to obtain the generic additional uncertainty for irregular fields in relative numbers this was normalized to \( s_{\text{air},0} \) for each photon beam

\[
\frac{s_{\text{air,irreg}}}{s_{\text{air},0}} = \frac{1}{10} \sum_{b=1}^{10} \left( \frac{\sigma_{\text{air}}(b)}{s_{\text{air},0}(b)} \right) - 2,
\]

where \( b \) represents the ten investigated photon beams.

The uncertainty estimations for the dose deposition \( q \) were calculated using the method presented by Nyholm et al.,\(^{20}\) which is based on a database including a large number of measurements in clinical photon beams. The uncertainties in the pencil kernels are separated into five beam quality intervals and described by a field shape component \( s_{q,A} \), which was created through comparison of the field shape characteristics of the treatment field and the reference field. This is then combined by square summation, i.e., assuming no correlation, with a depth dependent component \( s_{q,d} \) which is not related to the field setting \( A \). Hence, the predicted standard deviation for \( q \) was modeled as

\[
s_{q}(d;A) = \sqrt{s_{q,A}(d;A)^2 + s_{q,d}(d;A_{\text{ref}})^2}.
\]

Finally, the individually predicted standard deviations for \( \text{OF}_{\text{total}} \) were calculated as the square sum of the predictions of \( s_{\text{air}} \) and \( s_{q} \), assuming that these components also are independent from each other

\[
s_{\text{total}}(d;A) = \sqrt{s_{\text{air}}(d;A)^2 + s_{q}(d;A)^2}.
\]

### III. MATERIALS AND METHODS

All dose output values, both measured and calculated, were normalized to the isocenter point in the reference field size \( 10 \times 10 \text{ cm}^2 \). Besides the reference depth of 10 cm, both 5 and 20 cm were investigated using isocentric geometry when determining the total output factor (\( \text{OF}_{\text{total}} \)) in a large water phantom. Therefore the measured and calculated output values for the large water phantom do not only include effects associated with the size and shape of the fields, but also the depth has been included.

#### A. Linear accelerators, beams, and MLC types

Four different linear accelerators, each one equipped with an MLC associated with the respective manufacturer, were investigated. The Siemens Primus and the Varian Clinac 2300C/D were located at the University Hospital of Northern Sweden (Umeå, Sweden), while the Elekta SLi Precise and the General Electric Saturne 43 were located at the Department of Radiotherapy, Medical University Vienna (Austria). The nominal photon beam energies for the ten beams considered were in the range from 6 up to 25 MV. Table I summarizes the basic characteristics of the linear accelerators, photon beams, and MLCs used in the study.

#### B. Dosimetric equipment

At the University Hospital in Umeå all measurements were performed using a Scanditronix RK ionization chamber (cylindrical volume of 0.12 cm\(^3\)). When measuring \( \text{OF}_{\text{air}} \) the chamber was positioned inside two different semispherical build-up caps depending on the beam energy. The build-up cap for 6 MV photon beams was made of polymethyl methacrylate (PMMA) and the one for 18 and 20 MV of steel. Both caps were thick enough to exclude effects on the measurements related to electron contamination.\(^{21}\)

In all measurements performed at the Medical University Vienna a cylindrical ionization chamber (PTW type 31002, volume 0.125 cm\(^3\)) was used. During the measurements in
air a polystyrene miniphantom with a circular cross section of 3 cm in diameter and a measurement depth of 10 cm was utilized.

Measurements of $O_{\text{F}}$ air inside a miniphantom or a build-up cap describe in reality a ratio of two kerma values rather than a ratio of energy fluences, which is the definition of $O_{\text{F}}$ air in Eq. (1). However, as long as the difference in beam quality between the two fields is small the difference in the energy absorption coefficient $\mu_{\text{en}}/\rho$ will be negligible, and therefore the measured ratio will also reflect the ratio of energy fluence. The experimental results presented by Weber et al. indicate that this generally applies to the dosimetric equipment used for measurements of $O_{\text{F}}$ air in the present study. Nevertheless, for large field sizes in high energy photon beams they found small discrepancies, in the order of 0.5%, when comparing measurements performed with low-Z and high-Z build-up caps. We have therefore compensated for this effect in large field measurements utilizing the steel build-up cap, through small correction factors derived from the findings of Weber et al.

By considering the spread in equivalent measurements of $O_{\text{F}}$ air and $O_{\text{F}}$ total that were performed at different occasions we estimate that the experimental uncertainty, corresponding to one standard deviation of the measured values, was in the order of 0.2%. Moreover, when analyzing the correlation between observed deviations and predicted uncertainties one should bear in mind that this relation is not isolated from the precision in the experimental work. Hence, trying to evaluate the proposed method for estimating calculation uncertainties by a set of output measurements with a considerably different degree of precision, without any adjustments in the estimations, means that one should expect the validity of the predicted uncertainties to be affected.

### C. Irregular MLC fields

Six different field shapes were investigated. In order to test the capabilities of the calculation models thoroughly, the choice of shapes was based more on geometrical than clinical considerations. Furthermore, in two cases (the “circle” and the “belly”) three different sizes were included, which means that in total ten irregular MLC fields per photon beam have been studied (see Fig. 2).

These field shapes were used as templates in the local treatment planning system (Nucletron TMS, version 6.1) and the corresponding leaf settings were then created using the
midleaf intersection criteria. The leaf settings were exported to the linear accelerators via the Nucletron VISIR system (version 1.5) before irradiation. Finally, the same leaf settings were imported to the MATLAB 6.5 environment where the calculations were performed, from the TMS system through DICOM RT Plans. However, prior to the calculations for the Elekta accelerator, the leaf positions outside the field edges were altered in MATLAB since the DICOM RT Plan did not exactly describe the actual leaf setting on the accelerator during irradiation.

IV. RESULTS

The scatter plots in Fig. 3 display all output factors that were measured and calculated in this work and the corresponding deviations between calculations and measurements. Located in the upper right corner of each scatter plot is a histogram showing the analogous error distribution together with the associated mean value and standard deviation. The scale on the x-axis in Figs. 3(a) and 3(c) illustrates the wide range in dose output that has been investigated as a result of covering a depth span of 15 cm.

In Fig. 3(a) the fraction of calculated OF_total located within ±1% deviation from the measured value is 96%, while the largest discrepancy for OF_total amounts to +1.7% ("E" field shape at 20 cm depth in the 6 MV beam on the GE Saturne 43 machine). For OF_air [Fig. 3(b)] 97% of the 100 MLC fields can be found within ±1% deviation and the largest discrepancy amounts to −1.3% (large "belly" field shape in the 25 MV beam on the Elekta SLi machine). Figure 3(c) shows the deviations found for the dose deposition q for the 100 MLC fields and the three depths. The fraction of points located within ±1% deviation is 96%, while the largest discrepancy for q amounts to +1.5% (large "belly" field shape at 20 cm depth in the 6 MV beam on the Elekta SLi machine).

The outcome from Eq. (6), i.e., the average relative uncertainty introduced in OF_air by the irregular field shapes \( s_{air,irreg}/s_{air,0} \), was 1.6. Hence, in total the value for \( s_{air} \) was given by \( \sqrt{(1^2 + 1^2 + 1.6^2)} \), \( s_{air,0} = 2.14 \cdot s_{air,0} \). Following the expression for prediction of uncertainties in OF_total in Eq. (8), this estimation of \( s_{air} \) was then used as one of the components of \( s_{total} \).

Figure 4 contains a scatter plot of the observed deviations between calculations and measurements for OF_total (y axis) versus the corresponding predicted standard deviations \( s_{total} \) (x axis). As previously described the uncertainty is predicted individually for each point, hence the spread over the x axis. The dashed lines indicate 1, 2, and 3 predicted standard deviations, which means that in the case of a matching Gaussian distribution 68.3, 95.4, and 99.7% of the points should be included, respectively. The corresponding numbers for the sample of OF_total produced in this study are presented inside Fig. 4. The numbers follow the Gaussian behavior rather well, although the fraction included within one standard deviation is somewhat larger than expected. Additionally, in the upper left corner of Fig. 4 the corresponding distribution of normalized deviations, i.e., the ratio between the observed deviation (y value) and the predicted standard deviation (x value) for each point, is displayed. Theoretically, not considering statistical limitations related to the finite sample size, this should be a Gaussian distribution with a mean value of zero and a standard deviation equal to unity (indicated as a solid line in Fig. 4) if the uncertainty predictions are perfect in every case.

Another way of testing the association between the observed deviations and the predicted standard deviations is
deviations for OFtotal into 15 groups, based on the size of the
this study. This area was created by dividing all the 300
observed groups lie within the created 95% confidence interval, which sug-
the validity of the predicted uncertainties. Generally the mean values of the
predicted standard deviation
should constitute a Gaussian distribution with a mean value of zero and a
standard deviation equal to unity (indicated as a solid line).

displayed in Fig. 5. The area between the two lines can be
regarded as a continuous set of 95% confidence intervals for
a condensed data set representing the findings for OFtotal in
this study. This area was created by dividing all the 300
deviations for OFtotal into 15 groups, based on the size of the
predicted standard deviation s_{total} containing 20 data points
each. The average s_{total} (x value in plot) and the mean abso-
lute deviation (y value in plot) were then calculated for each
group (indicated as rings). For each one of those 15 groups
10 000 similar, but simulated, groups of 20 values were cre-
ated by random sampling from the corresponding Gaussian
distributions (using MATLAB 6.5), i.e., Gaussians defined by
the predicted s_{total} (x values of dots). The mean absolute de-
viation was then calculated for each group of 20 random
samples, resulting in a condensed distribution of 10 000
simulated absolute deviations (y values). Hence, all together
15 distributions of 10 000 simulated absolute deviations
were associated with the 15 original groups. The interval that
excludes the lowest 2.5% and highest 2.5% of those simul-
ated mean absolute deviations, i.e., corresponding to a con-
tinuous set of 95% confidence intervals, is located between
the two lines in Fig. 5. Practically all of the 15 groups of
observed deviations (i.e., the rings) lie within this area,
which means that the deviations that were found in this study
resemble typical samples from the predicted Gaussian uncer-
tainty distributions.

V. DISCUSSION

The actual deviations that are presented in Fig. 3 are gen-
erally very small. For instance, the total output factor
(OF_{total}) has an average absolute deviation of 0.4% and a
maximum deviation of 1.7%. This is lower than the devia-
tions found in a comparable investigation of a commercial
treatment planning system by Hansson et al., where the aver-
age value for the MLC fields (however not identical to this
set of field shapes) was 1.3% and the maximum deviation
amounted to 3.1%. Furthermore, in that case only the refer-
ce depth was studied for each beam (5 cm for 6 MV and
10 cm for 10 and 15 MV), as compared to the present work
where three depths were included for all ten photon beams.

The aim of the procedure for correcting the pencil kernel
characteristics is to remove systematic errors in the calcu-
lation of q that are associated with the size and shape of the
field as well as the calculation depth. The idea behind differ-
etiating this correction with respect to the radius is that it
should be applicable also to irregular field shapes, i.e., situa-
tions where the relative weight on dose contributions from
different radii in the field differs from the standard case of
squares where it originally was derived. Hence, the data set
presented here offers a good opportunity to perform a com-
prehensive test of this kernel correction. Figure 6 shows how
the deviations between calculations and measurements for q
are distributed when not applying the correction. There is a
noticeable shift toward positive deviations (overestimated
doses) in these calculations, particularly at larger depths.
This should be compared to Fig. 3(c) that shows the corre-
sponding distribution when employing the corrected pencil
kernels. In the latter case the distribution is much more con-
centrated around zero deviation, which also means that the
maximum deviation is reduced from 2.8 to 1.5%. The frac-
tion of calculated q that is included in the ±1% deviation
interval increases from 73 to 96% through the use of the
proposed pencil kernel correction.

The plots that are presented in Figs. 4 and 5 indicate that
the observed deviations for OF_{total} match quite well corre-
sponding Gaussian distributions, having individually pre-
dicted standard deviations for each situation. If on the other hand one examines the results in Fig. 3(b), describing observed deviations in $OF_{air}$, one can notice an unexpected group of negative deviations. After analyzing these underestimations in calculated $OF_{air}$, in more detail the conclusion is that the correlation between the input standard deviation ($s_{air,0}$) and the final deviations ($s_{air}$), which was present in the previous work by Olofsson et al., should be further analyzed. At least it can be questioned for the Elekta machine, which displays the lowest values for $s_{air,0}$, but the largest deviations for $OF_{air}$ in the end (all of the eight fields resulting in an underestimation of $OF_{air}$ between 0.8% and 1.3% are associated with the Elekta machine). This might be linked to the fact that the Elekta MLC is located closer to the source and the dose monitor chamber than the other collimators in the treatment head. This means that also the position of leaves located outside the open portion of the field will affect the values of $OF_{air}$ implying a risk for larger uncertainties.

One should also keep in mind that these samples of output factors are in fact not true random samples. The presented method for predicting uncertainties in output calculations is intended for a set of arbitrary fields in arbitrary megavoltage photon beams. Hence, repeating ten different field shapes in ten different beams could in fact mean that one encounters discrepancies from the estimated uncertainty not only due to the limited number of data points, i.e., random fluctuations, but also due to the fact that some of the field shapes and/or beams may be associated with small systematic errors in the output calculations. In that case increasing the number of field shapes, but not the set of photon beams, or vice versa, would not remove the reasons for invalid uncertainty predictions.

When discussing tools for independent verification of dose per MU one important issue concerns action levels, i.e., at what level of deviation should one start to question the treatment planning system? Without any knowledge or indication of the accuracy in the results coming from the verification tool this decision will in reality be a compromise between clinical tolerance limits and the generally unknown capacity of the verification calculation. However, this might also mean that the possibility of finding small errors that are related to flaws and weaknesses in the TPS is not utilized. Thus, finding the kind of errors that are perhaps not critical in each individual case, but that deteriorate the dosimetric precision and in the end will lead to poorer treatment outcome. A presentation of the inherent uncertainty together with the calculated output factors, as suggested here, can be a helpful tool when trying to catch such errors in the dose delivery. Nevertheless, it should be noted that the estimation of uncertainty that is presented here only is related to the calculation of dose per MU at the central beam axis in a homogeneous medium. Consequently, in order to include uncertainties associated with off-axis points or tissue heterogeneities in patients the models must be extended further. In particular the latter may constitute a considerable source of dosimetric uncertainty when treatment areas, such as the lung region, include substantial variations in density.

VI. CONCLUSIONS

In this work a model-based approach for calculating output factors in air and in water have been evaluated by comparisons with measurements at three treatment depths for a large number of irregular MLC fields and photon beams. We have shown that in spite of the fact that the calculation models were developed for a dose/MU verification tool using a minimum of commissioned input data, the results are very accurate. In a large majority of the investigated situations the deviations were smaller than 1% and in all cases smaller than 1.7%. Furthermore, for each individual setup the calculation model provides a prediction of the inherent uncertainty. A detailed analysis of the observed deviations suggests that these predicted uncertainties offer a valid and valuable estimation of the actual errors that can be expected when performing dose output calculations in homogeneous media. These uncertainty estimations can thus be used as a basis for creating action levels that may improve and simplify the clinical QA routines dealing with individual treatment plans.

ACKNOWLEDGMENTS

This work was partly supported by grants from the Cancer Research Foundation in Northern Sweden and partly by ESQUIRE within the framework of ESTRO.

Electronic mail: jorgen.olofsson@vll.se


T. Knöös, S. A. Johnsson, C. P. Ceberg, A. Tomaszewicz, and P. Nilsson,


