NEED FOR TREATMENT PLANNING PROGRAM VERIFICATION

James A. Deye, Ph.D. and Thomas N. Padikal, Ph.D.
Radiology Department
George Washington University Medical Center
Washington, D.C.

Frank E. Kearly, Ph.D.
Bureau of Radiological Health
Rockville, Maryland

Summary

Verification of results from the use of computers for treatment planning in radiation therapy is clearly an important concept, since a mis-applied treatment can be injurious to the patient. However, to date, there are no accepted testing procedures for verification of the accuracy of treatment planning computer hardware and software. This paper points to the need for such standardized verification procedures and suggests the criteria to be applied to any such efforts.

Introduction

Computers are being used routinely for the planning and calculation of patient dosage in radiation therapy centers throughout the world. This usage with its potential for increased efficiency and accuracy in the numerous calculations required by some treatment plans has aided the physician in his treatment decisions. However, there is also the potential for blind acceptance of the computed results and therefore mis-applied treatments.

Need for Verification

Applications at the leading edge of development such as those reported in this session by Dr. Bjarnard on computer controlled treatments or the complex calculational algorithm of Dr. Larson et. al. are being performed by only a few institutions and these are fully cognizant of their total responsibility. This is evidenced by the prominence each of these has given to verification procedures. Rather, it is those institutions which do not have adequate personnel, which are most likely to have computer errors which go unnoticed, that need guidance. And, it is the routine applications such as treatment planning that cause the greatest concern. This is due to the following aspects of computers applied to treatment planning.

a. Clinically important results: A mistake in a radiation therapy treatment plan can result directly in an incorrect dose to the patient, leading to either under-dosage and possible disease recurrence or over-dosage and/or possible unnecessary radiation damage to vital organs. Studies of clinical results have shown that for curative intent, a ± 5% difference in tumor dose is likely to be significant for many lesions.

b. Routine acceptance of the calculation methods: The essential algorithms of treatment planning have been established for some time now and are taken for granted. However, situations which exceed the original limitations of the programs may develop and go unnoticed.

c. Indirect interaction with the computer systems: Treatment planning systems are available commercially. Also, many third-party carriers pay for computerized results on a fee-for-service basis. These two facts may lead to situation where a user has a less than minimal knowledge and concern about the limitations or accuracy of the results.

A recent survey of users of one commercial
system found a fairly large variance (>5%) among institutions when each was requested to perform the same dosage calculation. These institutions were also employing quite different sets of data to describe the same treatment machine whose characteristics do not usually vary more than ±2%. An additional comparison was made by one of the authors (F.K.), using the data measured for 20 clinical and test situations and the results for these cases were computed on representative systems of the two most common radiotherapy treatment planning computers. Of 420 data points from these comparisons with measured data, 173 or 41% were outside ±5% agreement. The percentage was about the same from each system. Though it is not clear as yet how many of these discrepancies might be the result of measurement errors or how many might have been expected on the basis of known limitations of the software, the results certainly point to the need for further investigations of this type.

The need for verification of treatment planning programs is thus clear both on an "a priori" level as well as a demonstrable level. The questions remain as to who is responsible for such verification and what criteria should be employed in the validation procedure.

At the present time the responsibility for the validity of computer generated treatment plans rests primarily with the radiotherapy clinic which administers the treatment; however, it would be useful if commercial suppliers shared in this responsibility at least to the extent of supplying complete documentation for his system and cooperating with the user in the validation process. From the outset, it must be accepted that the validation of treatment planning systems is an ongoing process since the development and expansion of uses is continuous for such systems.

Suggested Criteria

Procedures and criteria for verifying computer generated treatment plans have not as yet been well established, though some attempts have been made. To date, the best efforts at such procedures are those of Saylor et al. who have produced a volume of measured data for various clinical situations to which one can compare computed results. In addition to these tests, and probably prior to using them, one should perform tests on the hardware of the system as well.

The overall principles to be kept in mind while formulating verification procedures are:

1. comparison of empirical data is the ultimate test,
2. tests should pertain to clinically relevant situations,
3. the data base, the algorithms and the hardware should be tested,
4. tests should be as diagnostic as possible, enabling the user to zero in on problems.

What are the criteria by which one judges the results of the verification procedures? First, the hardware should meet the manufacturer's specifications and be compatible with the constraints and capabilities of the software. Secondly, in the case of the software one should look for the mathematical correctness of the algorithms, which implies the availability of good documentation. Also, the more physical the algorithm is in its formulation, the easier it will be to recognize its limitations. Next, the criterion for any stored data base or algorithm constants is that they be well established either in the literature or from personal measurements, for their intended use in the program.

Finally, the criteria for accuracy of the comparison to empirical data should be realistic within a clinical framework. Calculations of relative doses or doses to non-critical structures might acceptably be less stringent than those involving tumor or critical structure doses so that experienced judgment must play a role in any set of verification criteria.

References


3. From the Treatment Planning Program Validation Project as distributed by W. Saylor, P. Heffron and T. Ames of the Div. of Radiation Therapy at the Univ. of North Carolina.
