Optimization in Brachytherapy

Gary A. Ezzell, Ph.D.
Mayo Clinic Scottsdale

mayo

Outline

- General concepts of optimization
- Classes of optimization techniques
- Concepts underlying some commonly available methods
- Specific brachytherapy applications
  - Permanent prostate implants
  - High dose rate brachytherapy

mayo

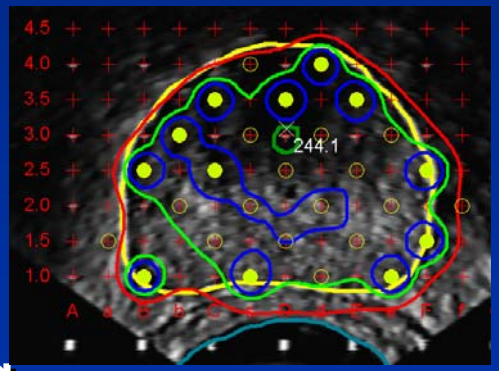
General concepts of optimization

- Big subject with huge literature
- This treatment will not be particularly mathematical
- Focus on key concepts and how they apply to brachytherapy
- The concepts also apply to IMRT optimization

mayo

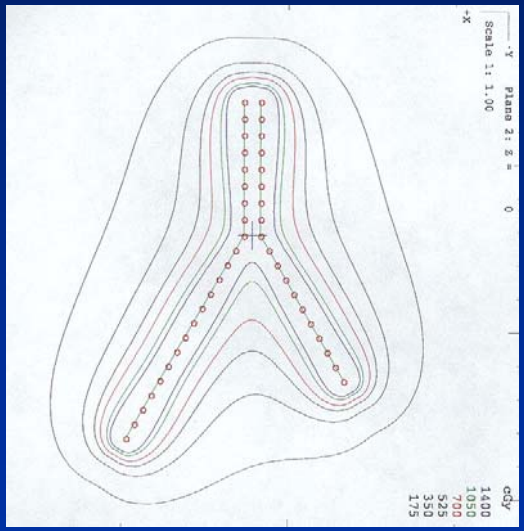
Examples of brachytherapy optimization

- Permanent prostate implants
  - Fixed seed activity
  - Choose seed locations to meet some objectives
    - target coverage
    - dose uniformity
    - rectal/urethral sparing



Examples of brachytherapy optimization

- HDR brachytherapy
  - Fixed activity, variable dwell times
  - Choose dwell positions and times
  - e.g. two-catheter endobronchial implant: sufficiently uniform dose at 1 cm from each of the catheters



Generalized brachytherapy problem

- Design a distribution of source terms such that the resultant dose distribution satisfies certain constraints and meets certain objectives as well as possible

mayo

Free variables

- Source locations, source strengths, and/or dwell times
- Depends on the specific problem
  - Prostate seed implant with template
  - HDR with implanted applicators
  - Stereotactic brain implants

mayo

Constraints and Feasibilty

- *Hard constraints*: cannot be violated
  - Physical (non-negative sources)
  - Clinical (cord dose < 45 Gy)
- *Soft constraints*: violation reduces plan quality (e.g. Rx dose to 98% of PTV)
- Any plan that satisfies the constraints is *feasible*

mayo

Feasible *vs.* Optimal

- Sometimes, a feasible solution is clinically acceptable
- More interesting problem: find a solution that *optimizes* some *objective*
  - *e.g.* "the dose to the surface of the prostate PTV is to match the prescription dose as closely as possible."

mayo

Express as a minimization problem

- Minimize the variance of the doses D_i at points i on the PTV surface from the prescription dose D_p

$$\text{Minimize } f = \sum (D_i - D_p)^2$$

- f is a simple *objective function*

Two competing objectives

- Prostate example:
  - Minimize the dose variation on the surface of the PTV
  - Minimize the dose to the adjacent rectum
- Cannot do both, must balance the two goals

mayo

Dealing with multiple objectives

- Most common approach: combine into a single objective function with *weighting factors* to control their relative influence

$$\text{Minimize } f = w_1 \sum (D_i - D_p)^2 + w_2 \sum (D_j - L_r, \text{ if } >0, \text{ 0 otherwise})$$

Uniformity on surface of PTV

Rectal doses below limit L_r

Dealing with multiple objectives

$$\text{Minimize } f = w_1 \sum (D_i - D_p)^2 + w_2 \sum (D_j - L_r, \text{ if } >0, \text{ 0 otherwise})$$

Dealing with multiple objectives

$$\text{Minimize } f = w_1\sum(D_i - D_p)^2 + w_2\sum(D_j - L_r, \text{ if } >0, 0 \text{ otherwise})$$

Additional objectives

- Add penalty based on number of needles

$$\text{Minimize } f = w_1\sum(D_i - D_p)^2 + w_2\sum(D_j - L_r, \text{ if } >0, 0 \text{ otherwise}) + w_3 N$$

How good is that function?

- The objective function is a mathematical model of the clinical goals.
- Does the model capture the essence of the clinical thinking?

$\approx$ Minimize $f = w_1\sum(D_i - D_p)^2 + w_2\sum(D_j - L_r$, if >0, 0 otherwise) ?

How to steer the results?

- What are the optimization parameters?

$$\text{Minimize } f = w_1 \sum (D_i - D_p)^2 + w_2 \sum (D_j - L_r, \text{ if } >0, \text{ 0 otherwise})$$

How to steer the results?

- Aren't these chosen by the physician?

$$\text{Minimize } f = w_1\sum(D_i - D_p)^2 + w_2\sum(D_j - L_r, \text{ if } >0,\ 0 \text{ otherwise})$$

How to steer the results?

- Aren't these chosen by the physician?

$$\text{Minimize } f = w_1\sum(D_i - D_p)^2 + w_2\sum(D_j - L_r, \text{ if } >0,\ 0 \text{ otherwise})$$

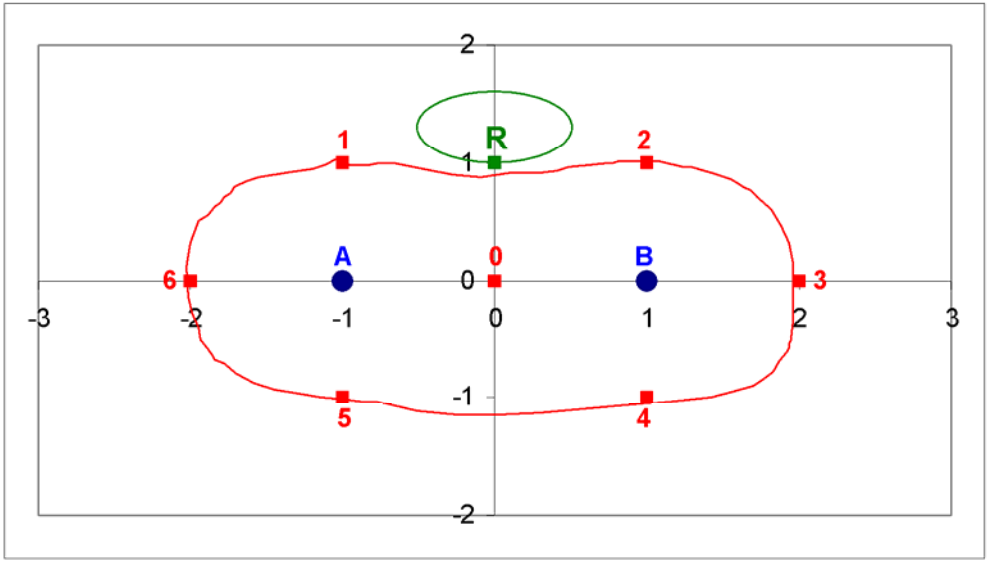
- Yes and no.
  - It may be that the doses the physician really wants are not the best to use as input

Three uses of the word “prescription”

- “Prescription”
  - The dose objectives you want to achieve
  - The dose parameters you give to the optimizer
  - The dose results you eventually accept
- Ideally, all are the same. Not always the case

mayo

An idealized problem



$D_0 = A + B$

$D_1 = D_5 = A + B/5$

$D_2 = D_4 = A/5 + B$

$D_3 = A/9 + B$

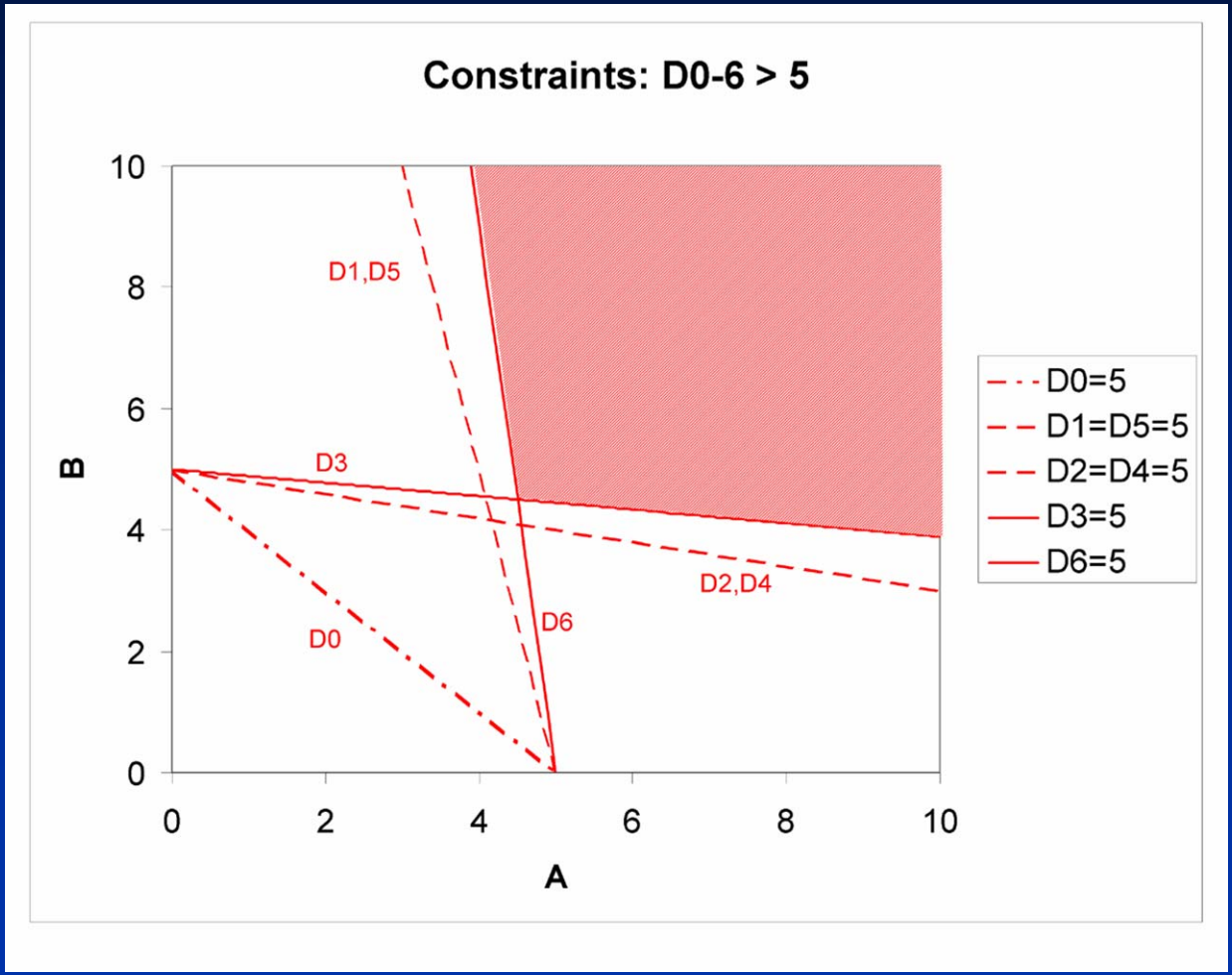
$D_6 = A + B/9$

$D_R = A/2 + B/2$

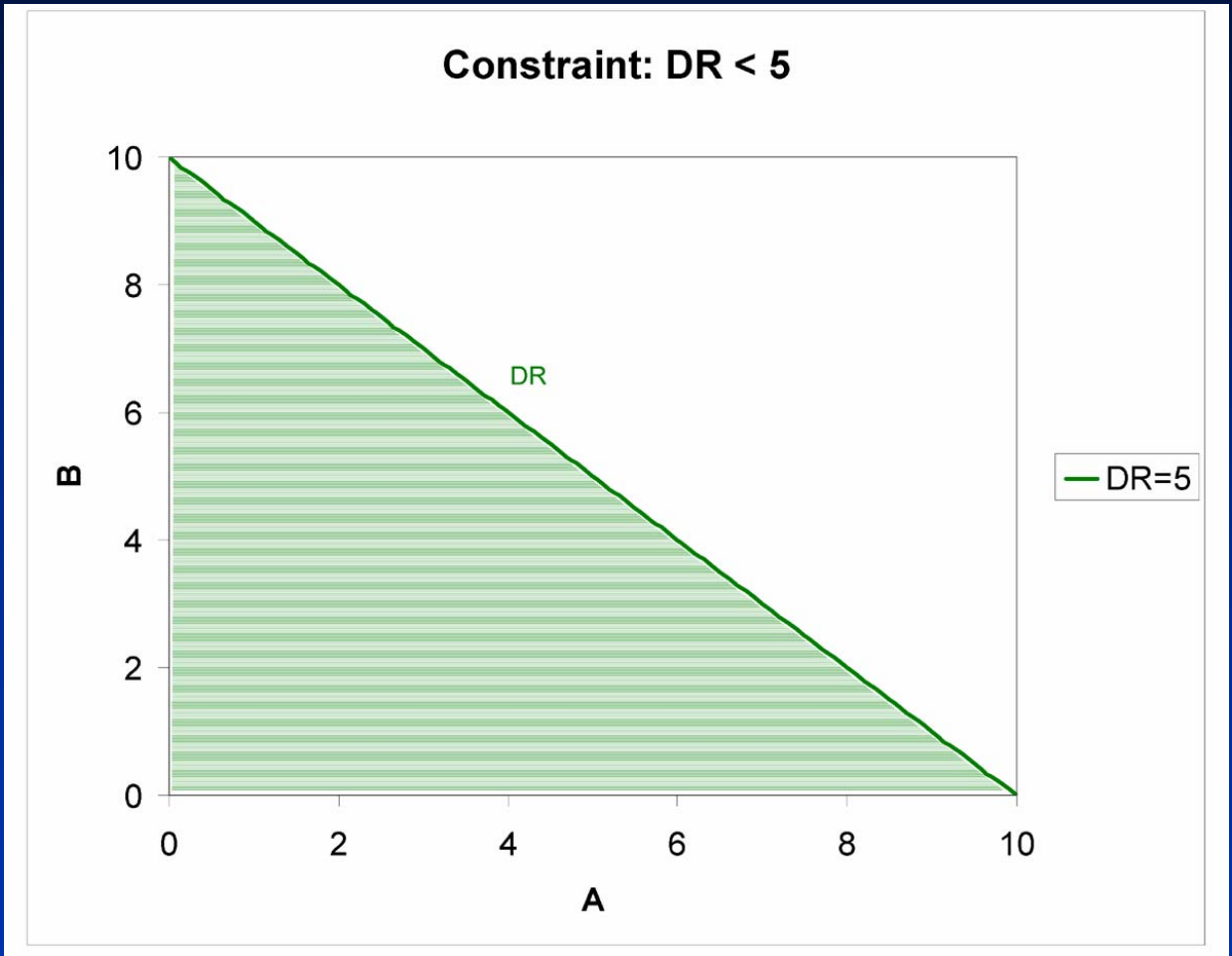
"Dose" depends only on distance, no attenuation …

1 source strength unit at unit distance → 1 dose unit

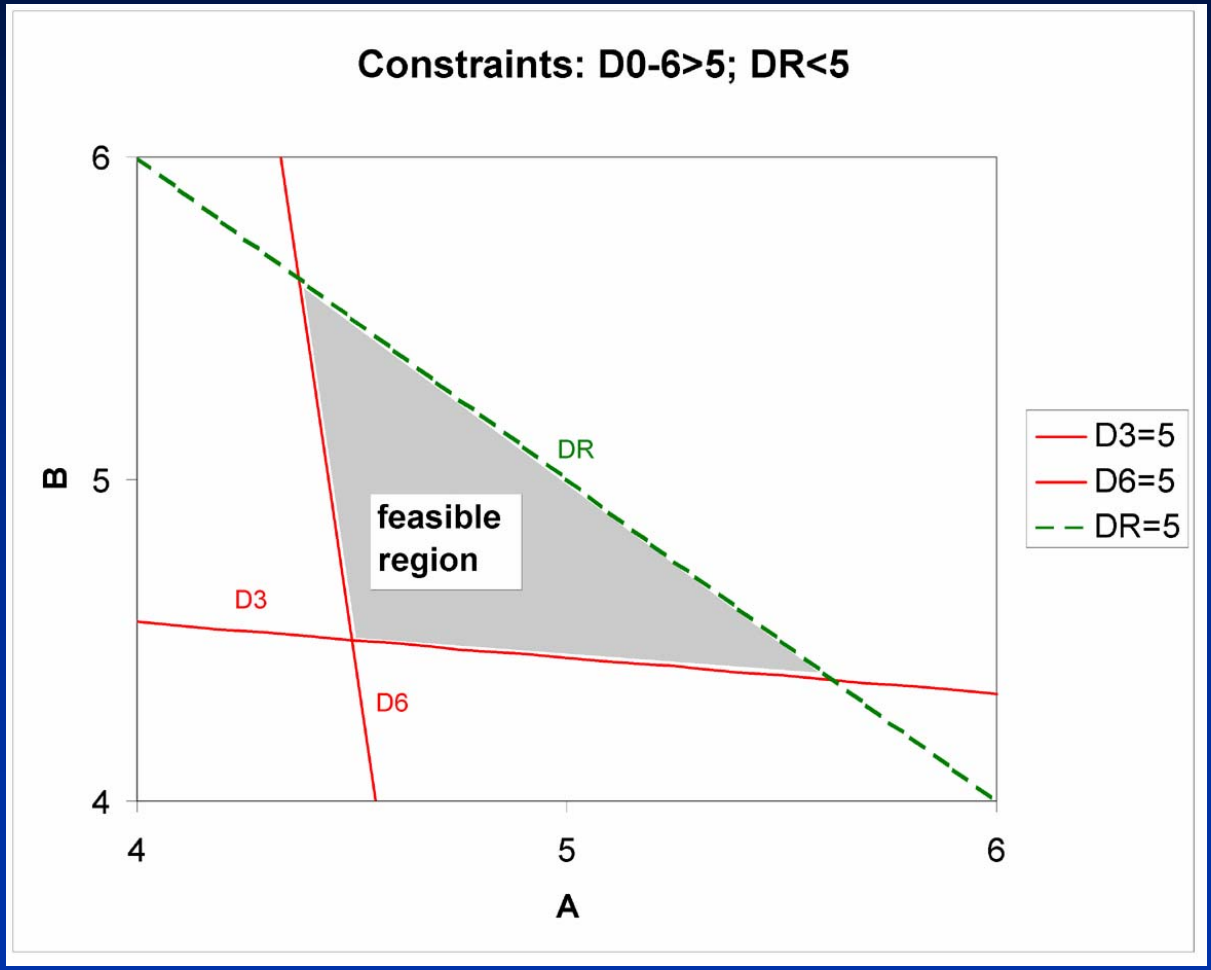
Target is to receive at least 5 dose units



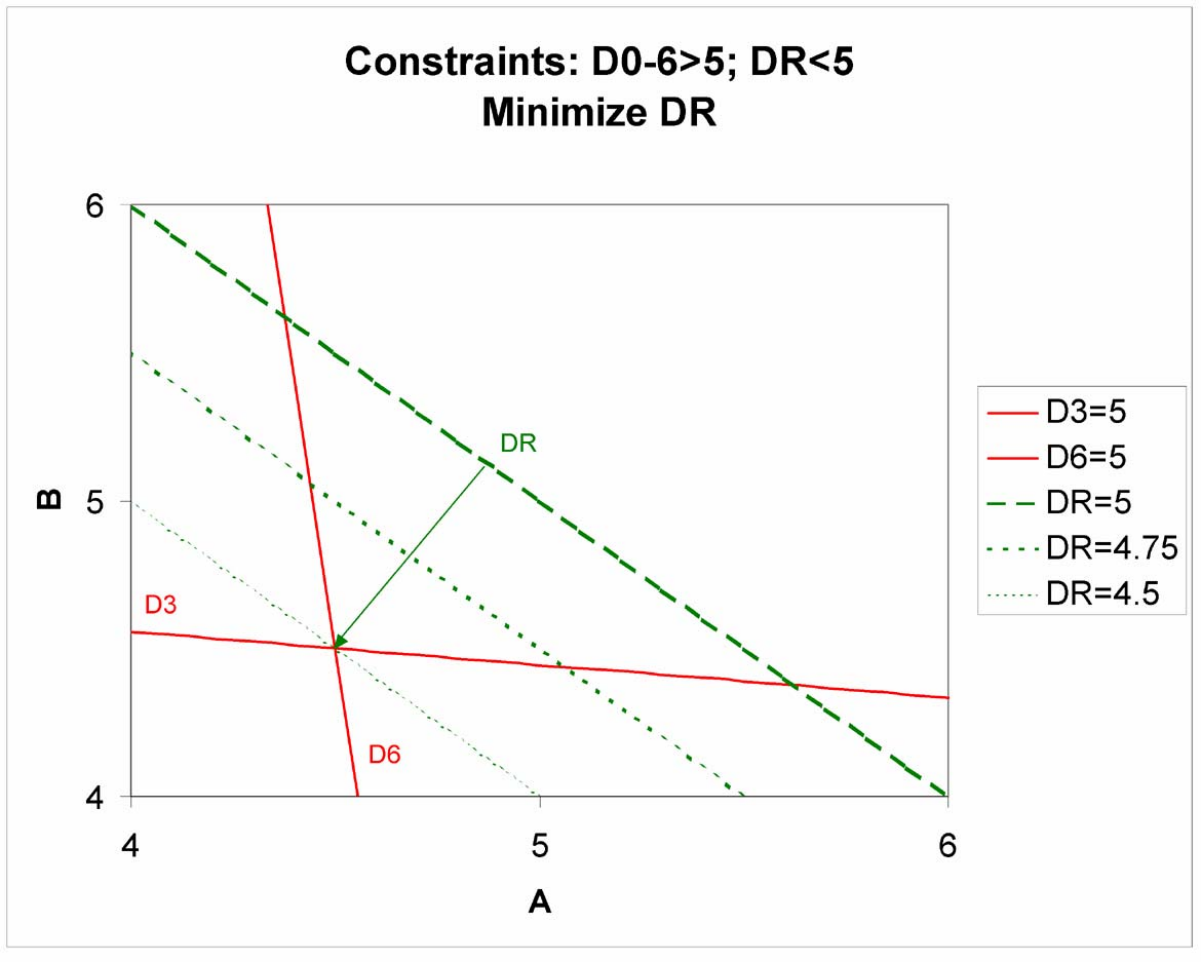
Rectum is not to exceed 5 units of dose



Feasible solutions for both constraints



Optimize dose to rectum



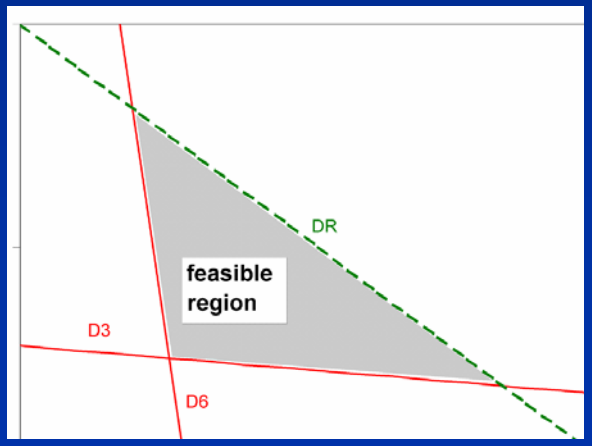
Optimum:

A=B=4.5

Features of linear systems

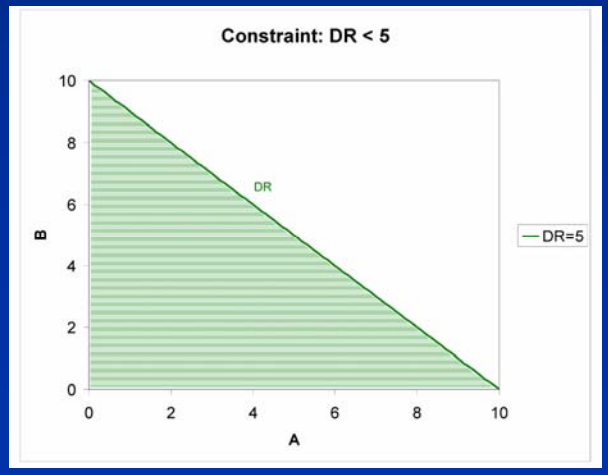
- Feasible solutions to sets of linear constraints are bounded by a *convex* multidimensional polyhedron. (This simple example is two-dimensional.)

Convexity means that you can move from one feasible solution to another along a straight vector in the space without ever leaving the feasible region.



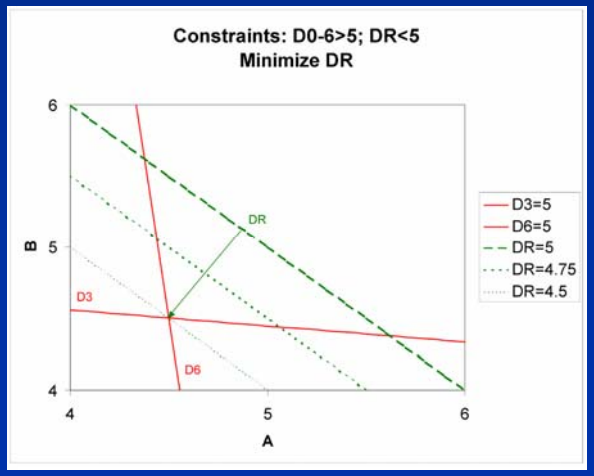
Features of linear systems

- Inequalities define regions of the space. Equalities define surfaces of lower dimensionality. (DR<5 defines a section of a plane, while DR=5 defines a line.)



Features of linear systems

- When an optimum solution is sought by maximizing or minimizing a linear function of the variables, the optimal solution will lie on a vertex.

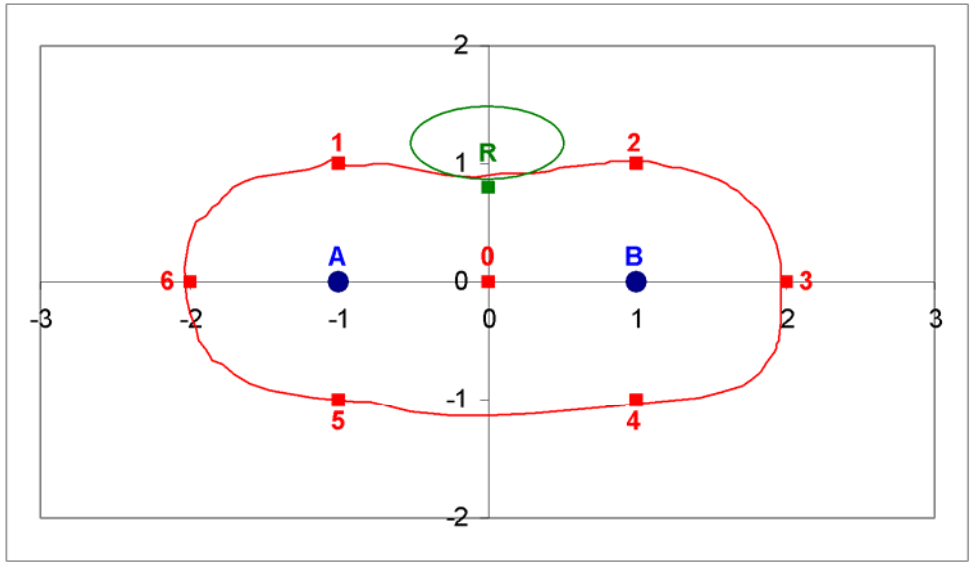


Quadratic functions

- For quadratic functions, the surfaces are curved, but similar concepts apply

More interesting problem

- Move rectum closer; point R now at 0.8 units from origin



Competing objectives

$$T = \left(\sqrt{\sum_{i=1}^{6} (Di - 5)^2} \right) / 6$$

Average deviation of target dose points from prescription “5”

$$S = D_R = A/1.64 + B/1.64$$

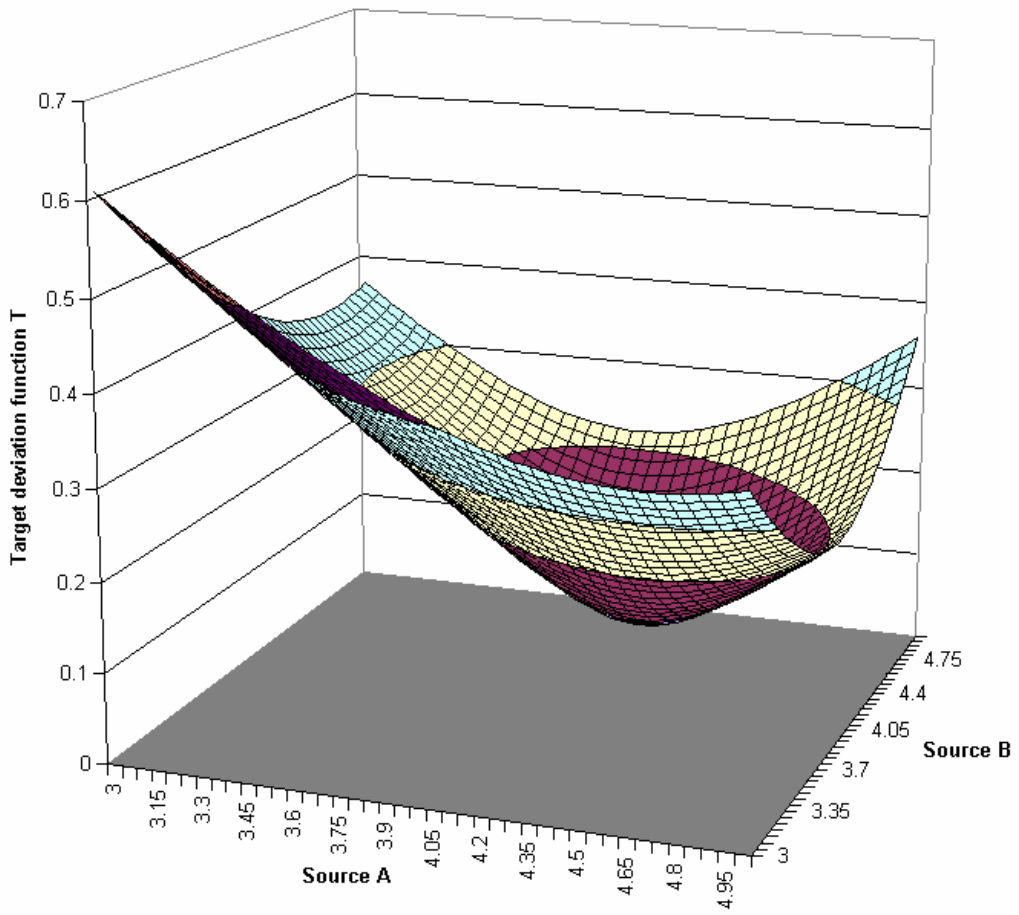
Rectal dose point

Combine in an objective function with normalized weighting factors

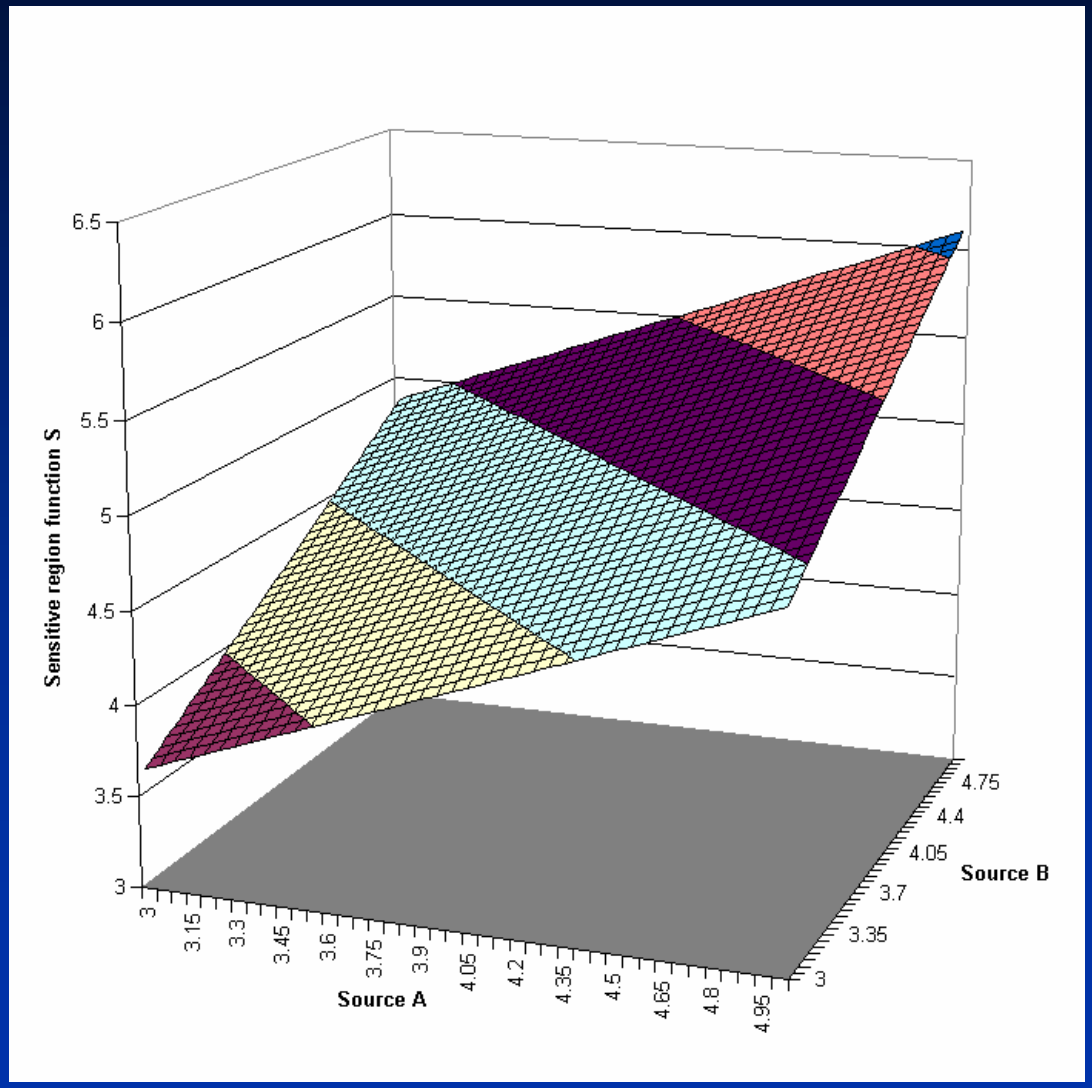
$$F = (w_1 T + w_2 S) / (w_1 + w_2)$$

mayo

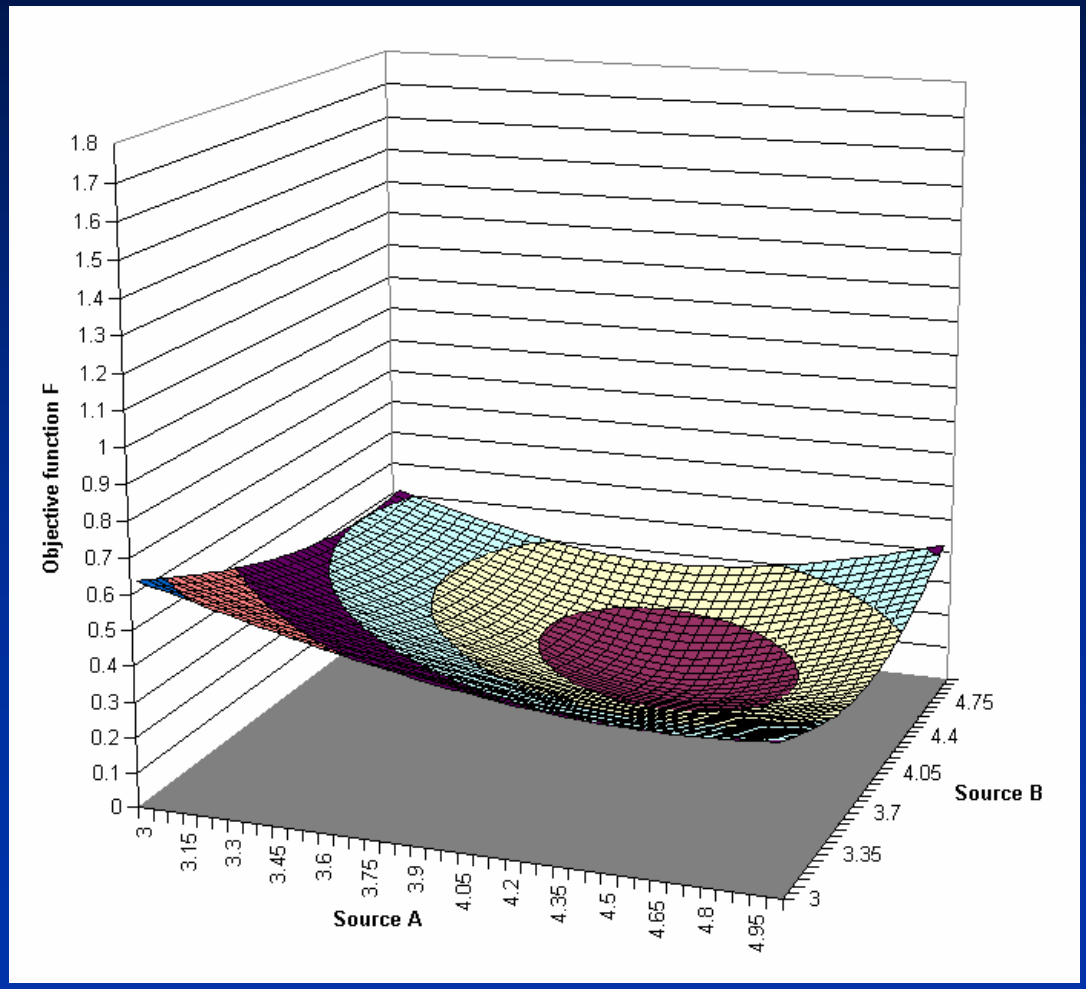
Variation of Target component



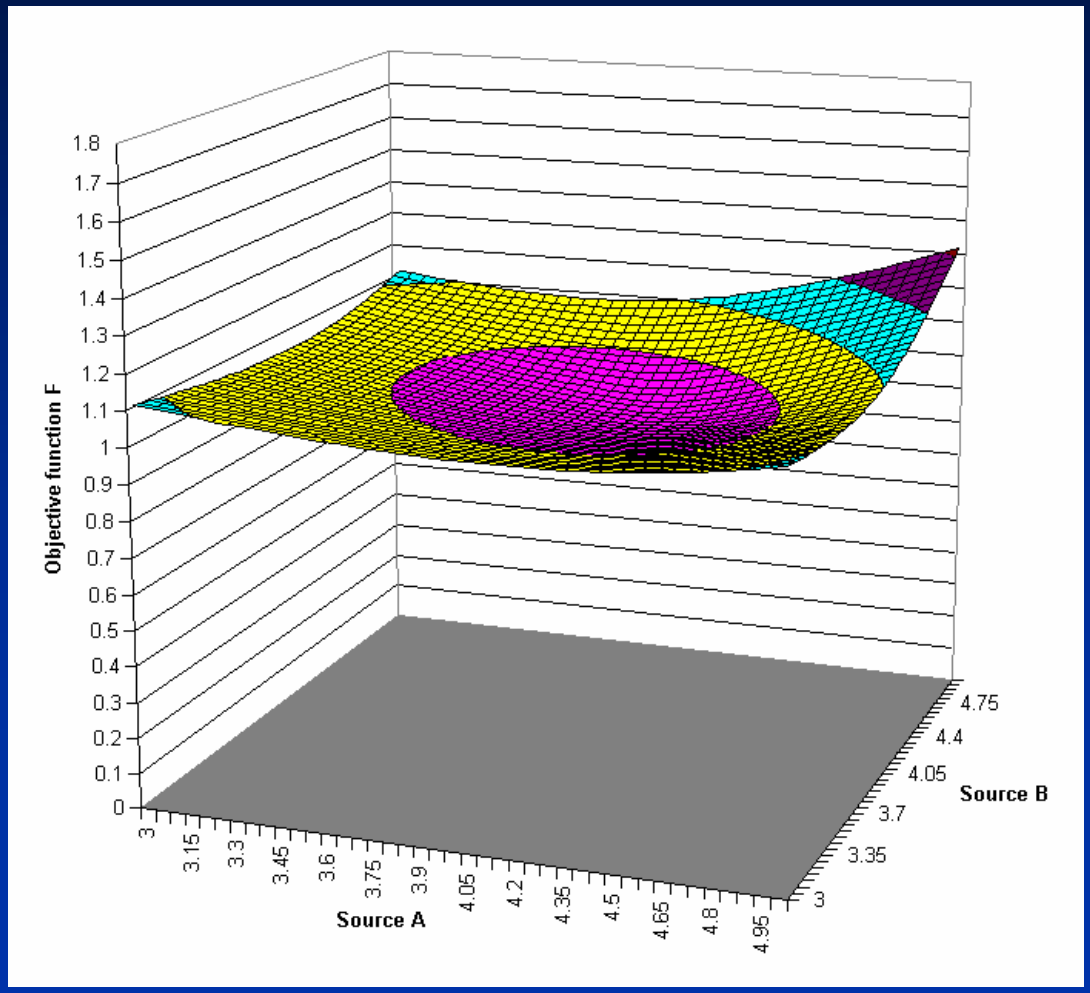
Variation of Structure component



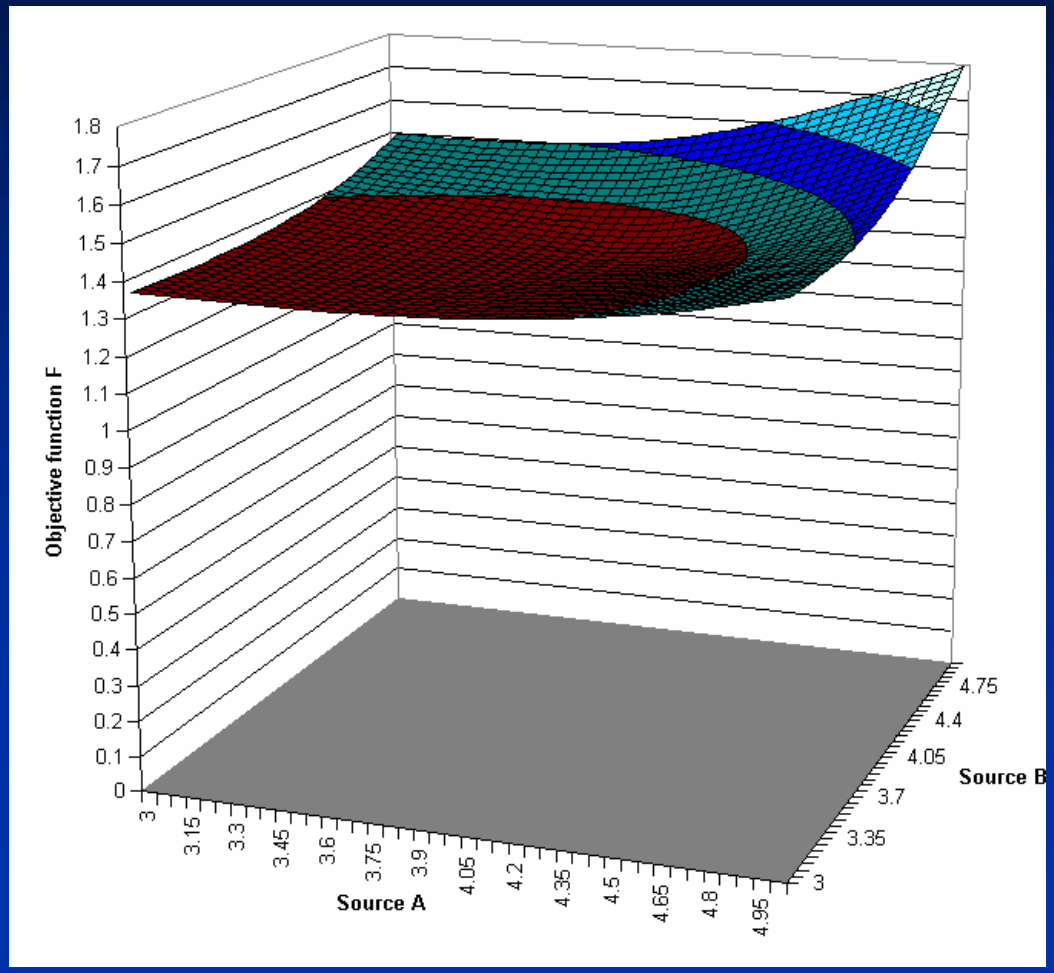
Solution space for w_2=100, w_1=1



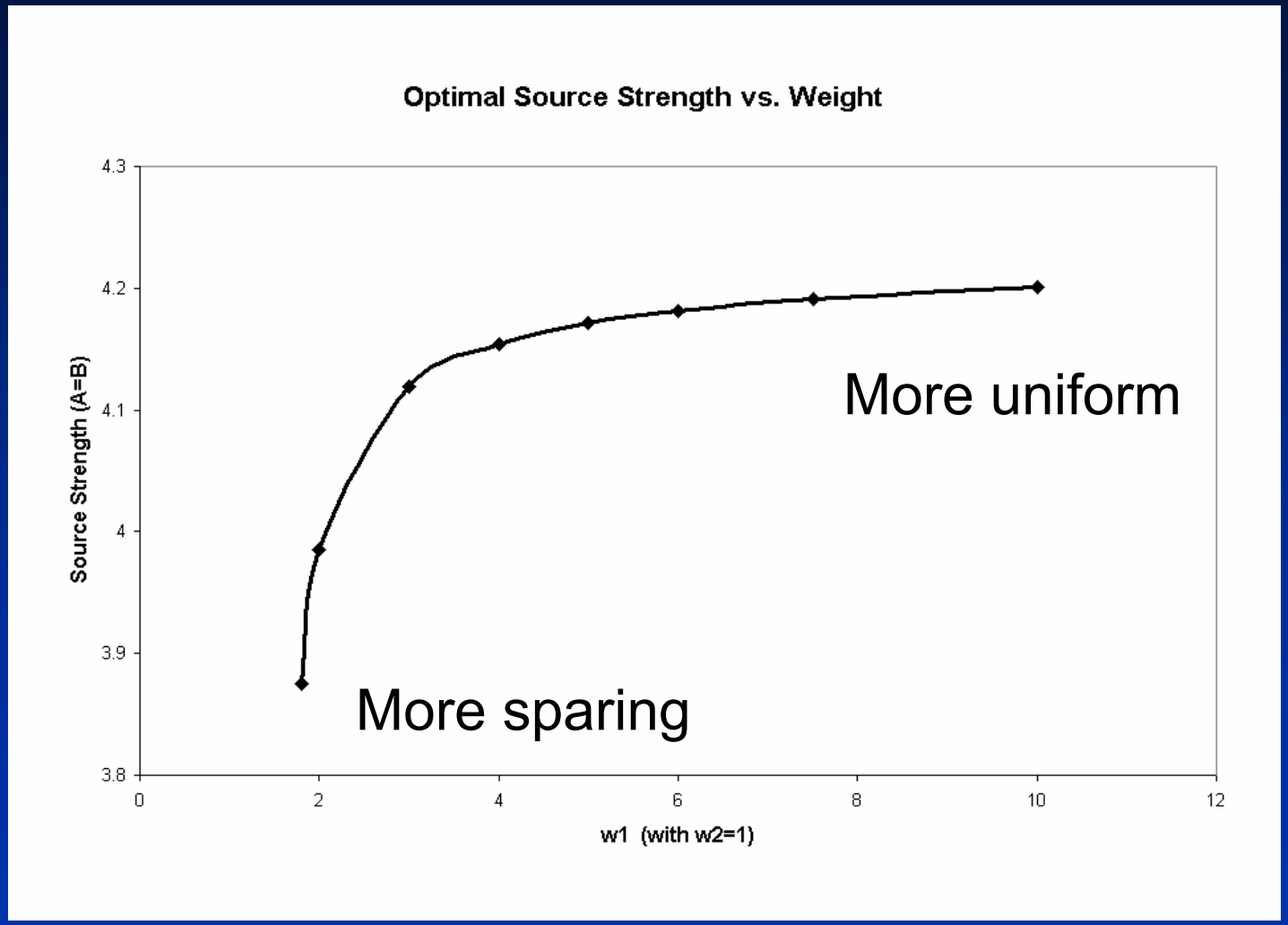
Solution space for w_2=5, w_1=1



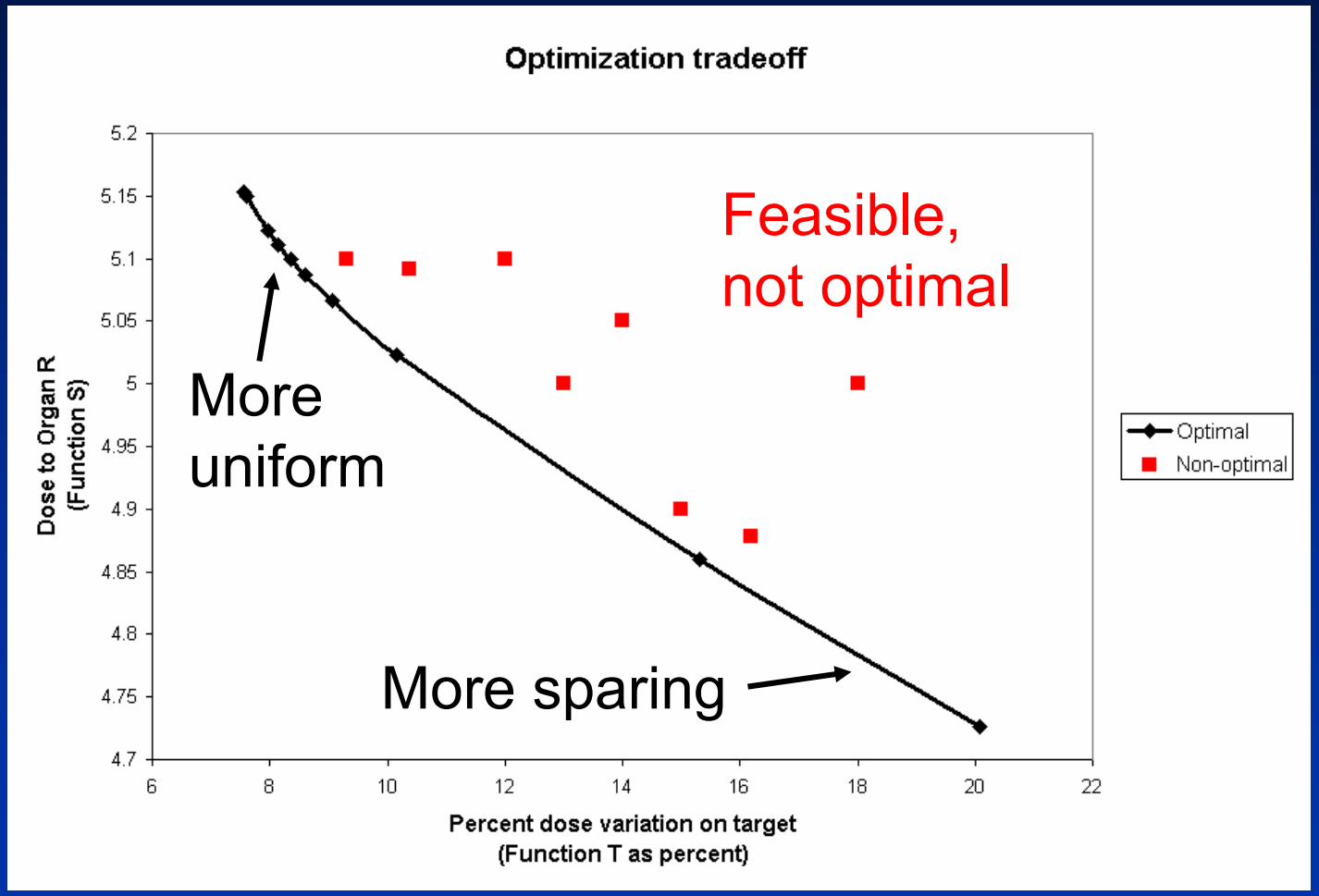
Solution space for w_2=3, w_1=1



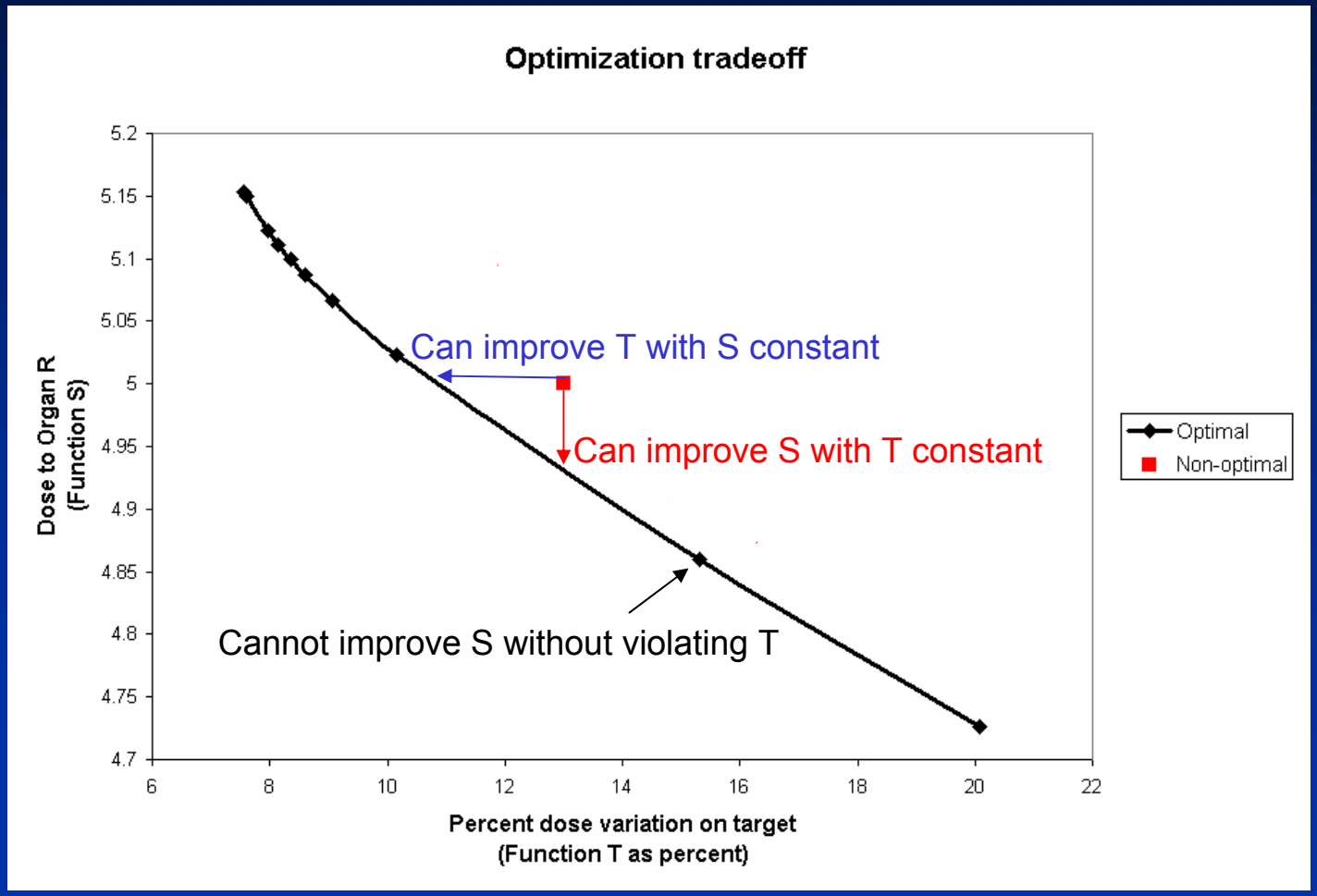
Location of optimum shifts with weights



Optimization tradeoff: *Pareto front*



Optimality: no objective can be improved without worsening another one



Practical example of Pareto front

- If a physician wants to cover a prostate while sparing the rectum, the planner can provide a series of plans, with, for example, 98%, 95%, 93%, and 90% target coverage, each with the best obtainable rectal dose.
- Such a series of results represents the Pareto front for the problem at hand and shows the best available choices.

mayo

General approach to optimization

- Decide which objective is most important
- Decide the limit, or range of limits, to which that objective may be pushed
- Optimize the other objectives by pushing the primary objective to this limit

mayo

Related issue: the evaluation problem

- The objective function is not available to the user, only the “steering parameters”
- The function may not encode the quality measure preferred by the planner: e.g. DVH, conformity index, TCP, EUD, …

mayo

“Evaluation problem”

- Carefully think through how to quantitatively evaluate the results
- Use the mechanics of the optimizer software to control those results
- Cannot assume that what the computer declares is mathematically optimal according to its algorithm is actually clinically optimal

mayo

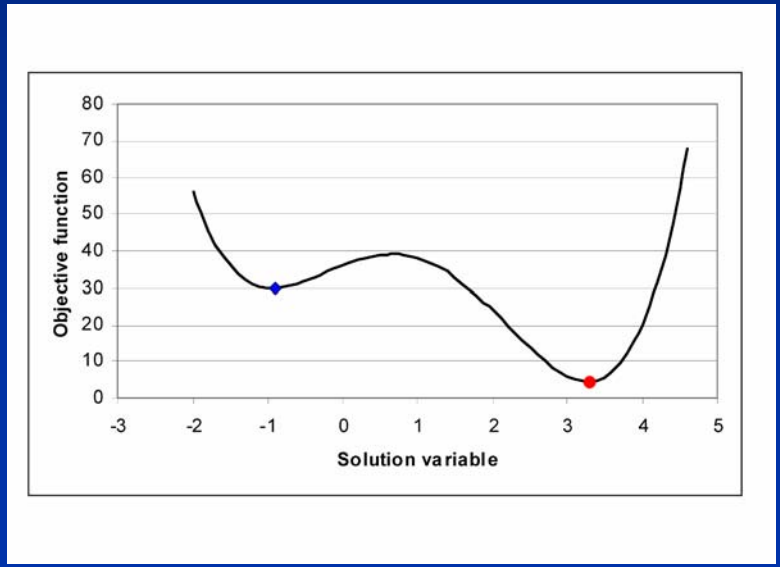
Robustness problem

- When planning an implant, need to think about the effect of uncertainty in carrying it out
- e.g. for prostate: A plan with more seeds of somewhat lower activity may be more stable in the face of uncertainty than one with fewer seeds of higher activity

mayo

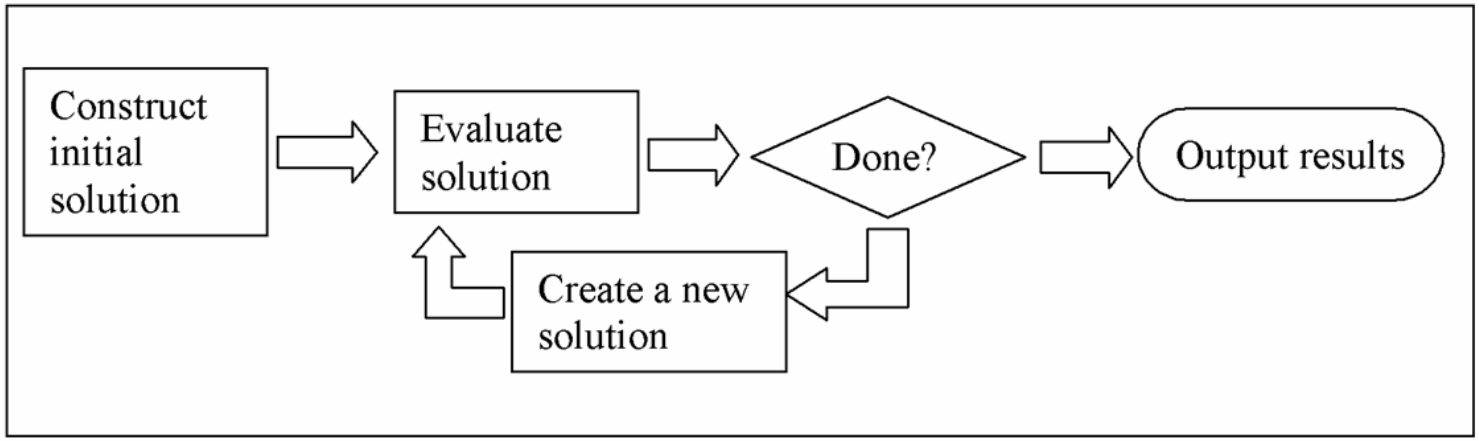
Local vs global optimum

- Some objective functions can have multiple minima (e.g. with dose/volume constraints; multiple possible source positions)
- Some search process can escape from local minima, some cannot (more later)



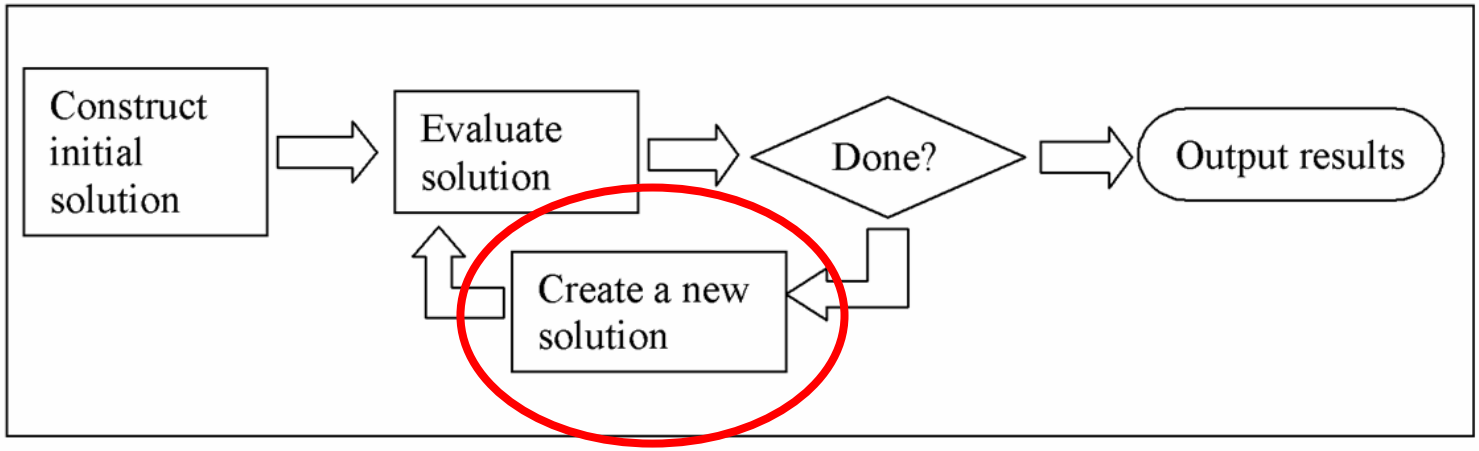
Classes of optimization techniques

- Most approaches are iterative



Classes of optimization techniques

- Most approaches are iterative



- Differ in how to create new solutions

Classes of optimization techniques

- Deterministic
- Stochastic (probabilistic)
- Deductive
- Heuristic, or phenomenological

mayo

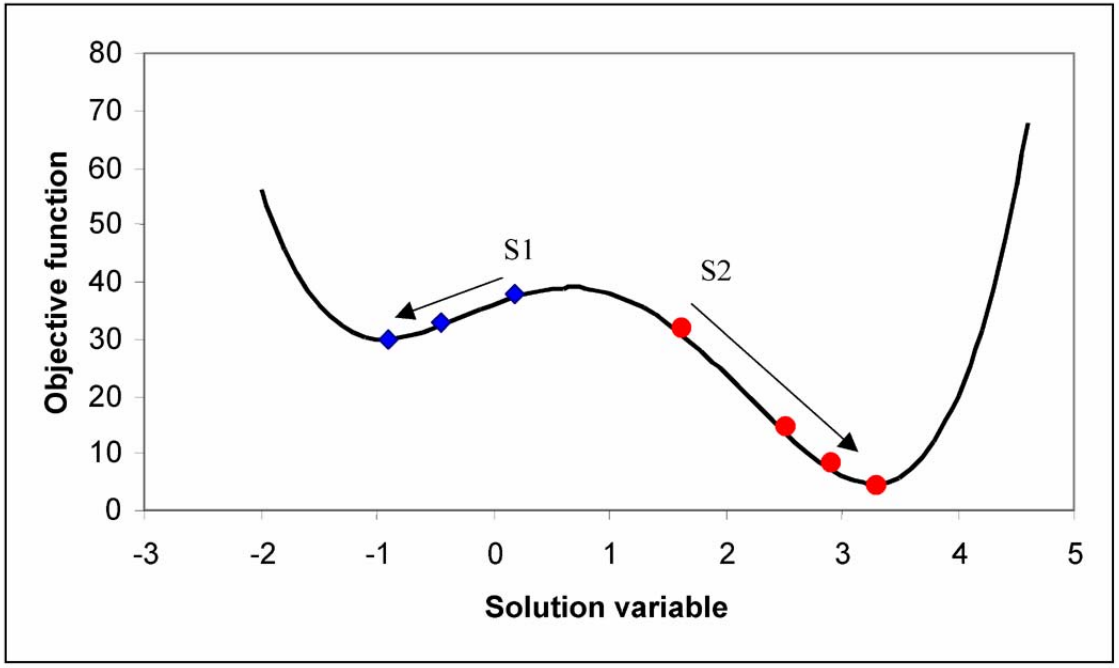
Deterministic methods

- Move “downhill” from the starting point according to an algorithm (steepest descent, conjugate gradient, Nelder-Mead simplex, …)
- May use calculations of gradient in solution space

mayo

Deterministic methods

- Fast, but cannot escape from local minima (may not be a problem!)



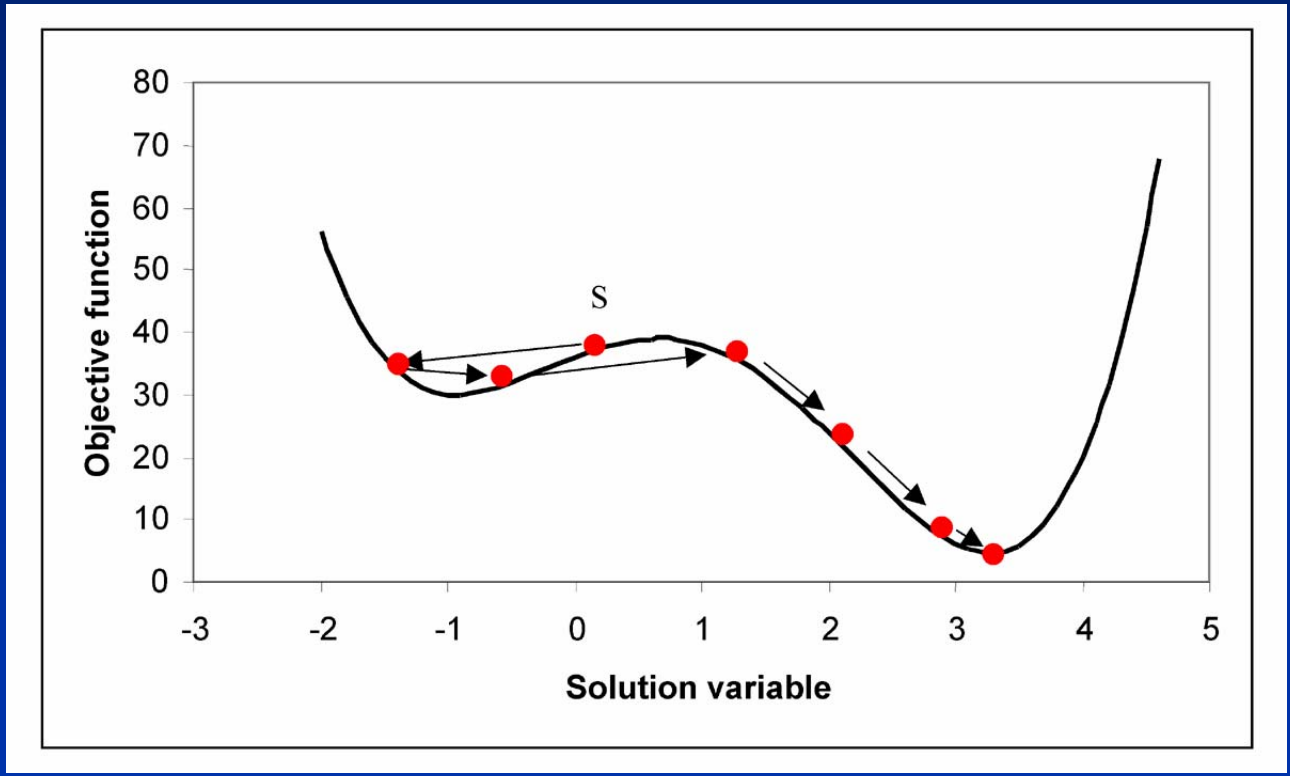
Stochastic (probabilistic) methods

- Use randomness in the search process (simulated annealing, genetic algorithms)
- Can be applied to objective functions of any mathematical form

mayo

Stochastic methods

- Slower (but may not be a problem!)
- Can escape from local minima



Deductive methods

- Test solutions in a systematic, logical sequence that is designed to eliminate subsets of potential solutions from consideration (e.g. branch and bound)
- Can be applied to objective functions of any mathematical form

mayo

Heuristic methods

- Use related relationships to arrive at an adequate solution to the original problem (e.g. geometric optimization)
- Related relationships sometimes are called “adjoint functions”
- Can be very fast and produce results that are useful without being “optimal”

mayo

More on two probabilistic methods

Simulated annealing

- Analogous to a crystal seeking its lowest energy state as it slowly cools from an initially high temperature (Metropolis 1953)
- Early application to brachytherapy by Sloboda (1992)
- At least one current commercial application: VariSeed

mayo

Simulated annealing algorithm

1. Choose initial solution and evaluate
2. Create another in by taking a “step in solution space” in a random direction

   The size of the step depends on a “temperature” parameter, initially large.

   The temperature is reduced as the iteration proceeds

mayo

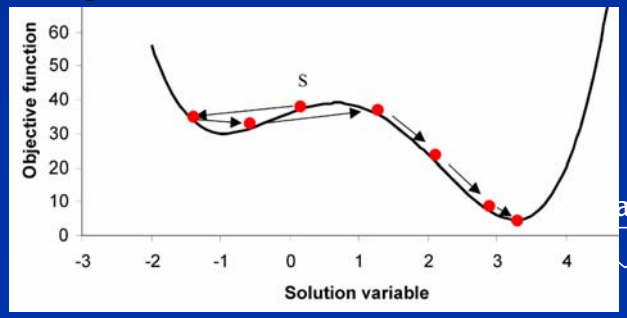
Simulated annealing algorithm

3. Evaluate the new solution and compare to previous (ΔF). Accept if better. If worse, still accept with a probability that depends on the temperature and the size of the deterioration

$$p = e^{-\Delta F / kT}$$

Simulated annealing algorithm

4. Iterate, using a “cooling schedule”, thus shortening the steps and reducing the likelihood of uphill moves.
5. By reducing the temperature sufficiently slowly, the system is allowed find the region of solution space that contains the global optimum and converge to it



Simulated annealing algorithm

- Slow (logarithmic) cooling can guarantee global optimality
- Faster cooling schedules and different probability distribution forms speed the convergence

mayo

Genetic algorithms

- Operate on populations of alternative solutions
- Mimic processes of evolution
  - Mutation
  - Mating with crossover of genes
  - Replication in the next generation proportional to fitness

mayo

Genetic algorithms

1. Construct a population of individual solutions, each being a string of numbers encoding the values of the free variables. Typically, each solution is a bit string of concatenated binary numbers

e.g. for prostate implant: “1” means seed position is “on”; “0” means is “off”

`101011010011010011010101110100 1001`

Genetic algorithms

2. Evaluate each individual using the objective function, and then assigning a fitness value to each
3. Produce a new population, or generation, from the old using a combination of operators

mayo

Genetic operators

a. Proportional replication: select a string for representation in the new generation with a probability dependent on its fitness
b. Crossover: pairs of selected strings “mate” and produce offspring that contain parts of each parent
c. Mutation: introduce variation by altering randomly selected bits in some of the strings

mayo

Crossover

1010110100110100110101011101001001

0110100101011010010111111010110000

1010110100110100010111111010110000

Profitable patterns are perpetuated

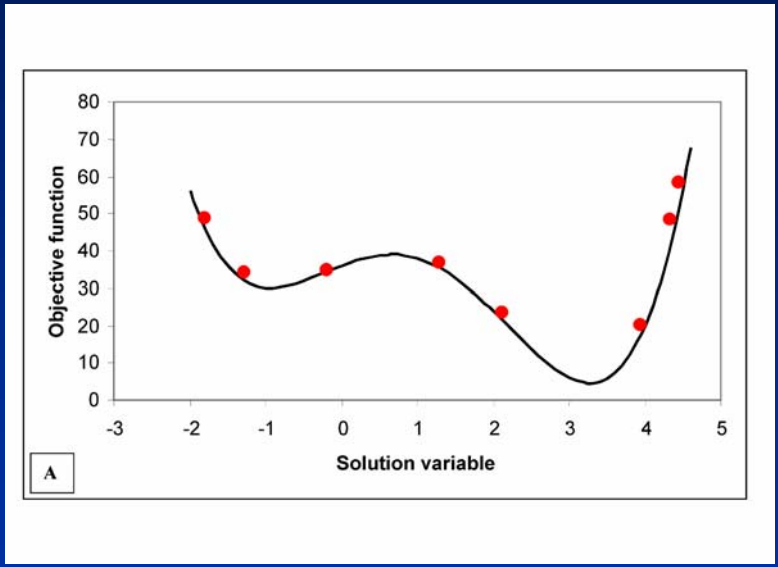
Best solutions “mate” most frequently

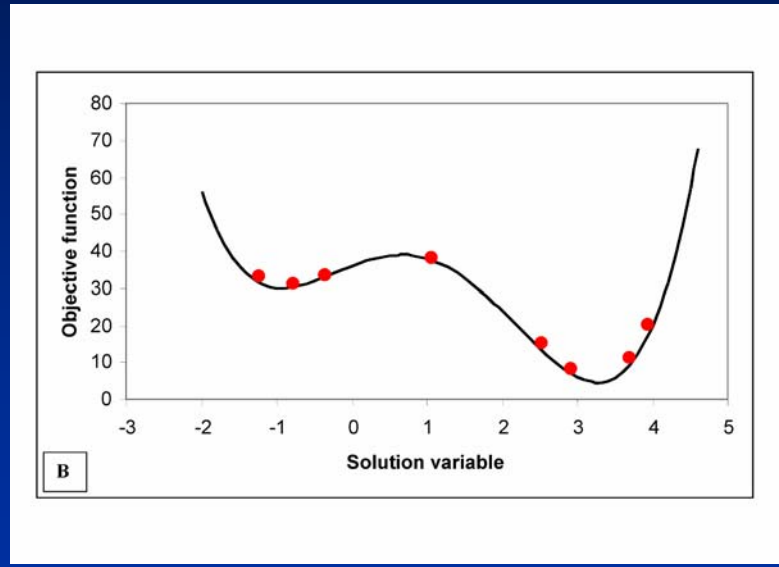
1010110100110100110101011101001001	quarterback
0110100101011010010111111010110000	cheerleader

1010110100110100010111111010110000	medical physicist

Profitable patterns are improved

Genetic algorithms: population converges to good solutions





Yu and Reinstein (1996) described a genetic algorithm that was later developed into clinical tool, Prostate Implant Planning Engine for Radiotherapy, PIPER, (Yu et al. 1999) and used for intraoperative planning (Messing et al. 1999).

Geometric Optimization

- Heuristic method developed by Edmundson (1993) for HDR dwell times
- Assume that sources are distributed throughout the implant volume (prostate, breast, …)
- Want dose uniformity between the sources
- Do not want to define calculation points between the sources

mayo

Geometric Optimization

- Look at source locations themselves
- Want the dose to each location from all the other sources to be uniform
- Set dwell time for each to be inversely proportional to the sum of the inverse square of the distances to the other sources

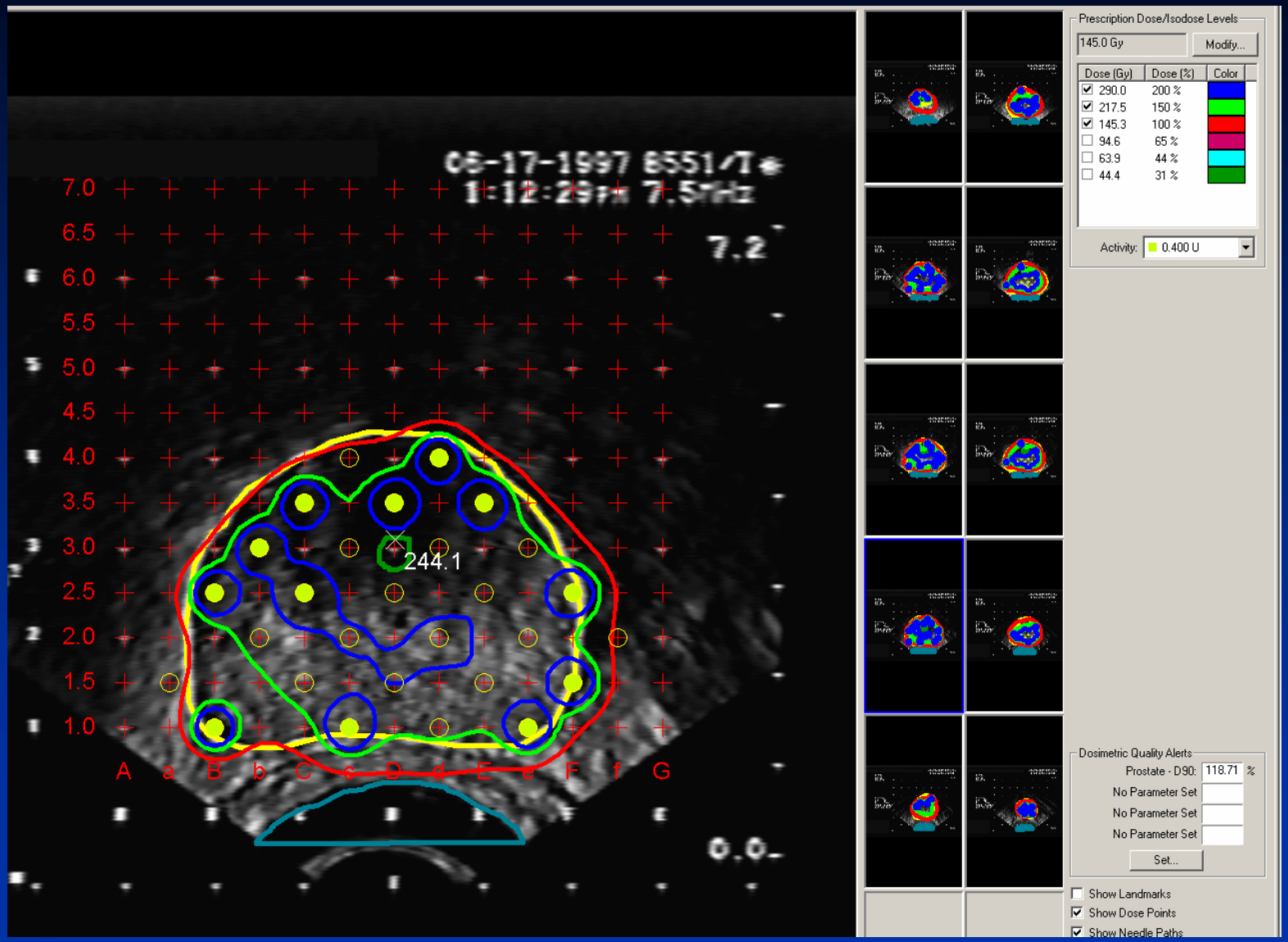
$$T_i \propto \left[\sum_{j=1,\, j \neq i}^{Sources} \frac{1}{r_{ij}^2} \right]^{-1}$$

Commercial systems

- Nucletron
  - Dose point optimization (analytical solution by singular value decomposition)
  - Geometric optimization
- BrachyVision
  - Dose optimization (downhill search by Nelder-Mead simplex)
  - Geometric optimization
- VariSeed
  - Simulated annealing

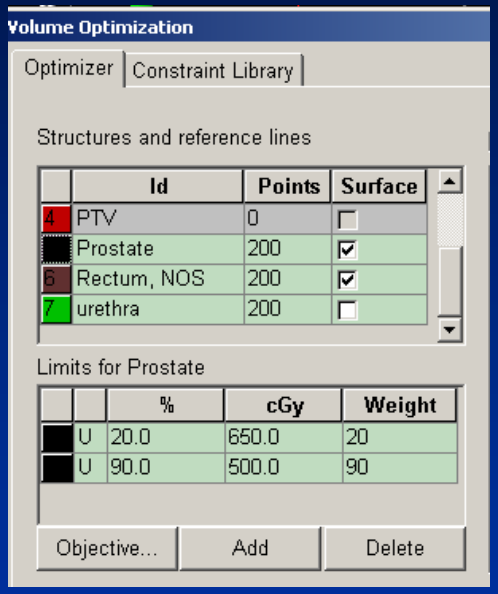
mayo

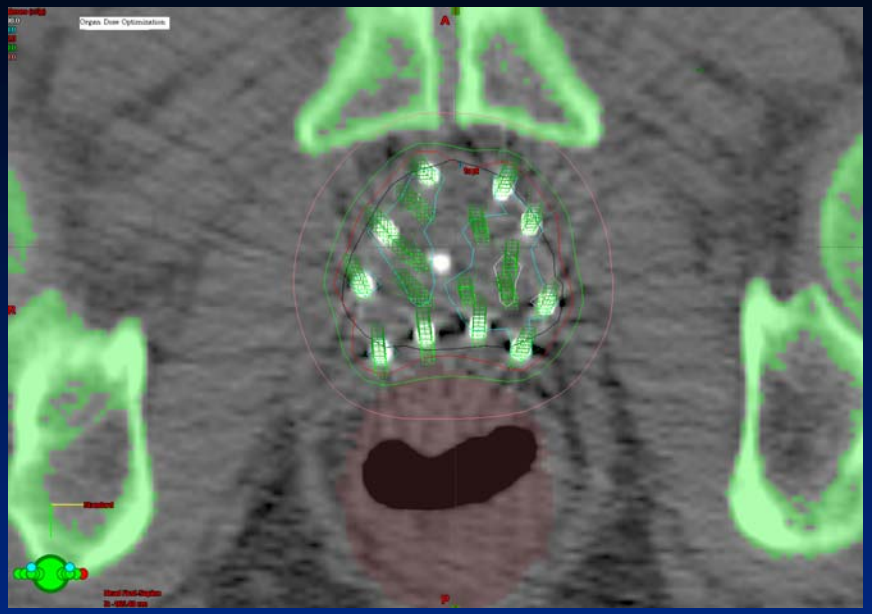
Intraoperative planning

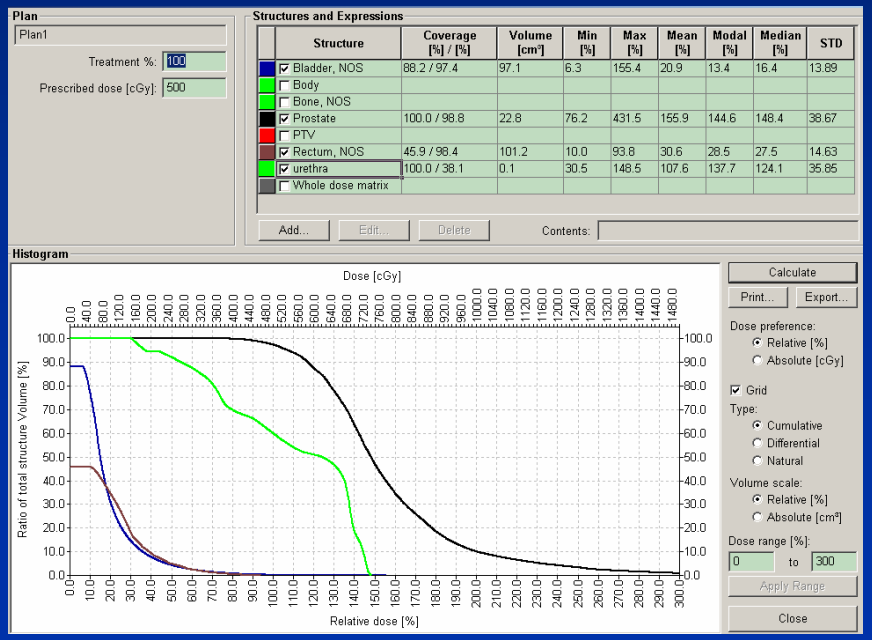


Courtesy Michael Mariscal, Varian

Dose optimization

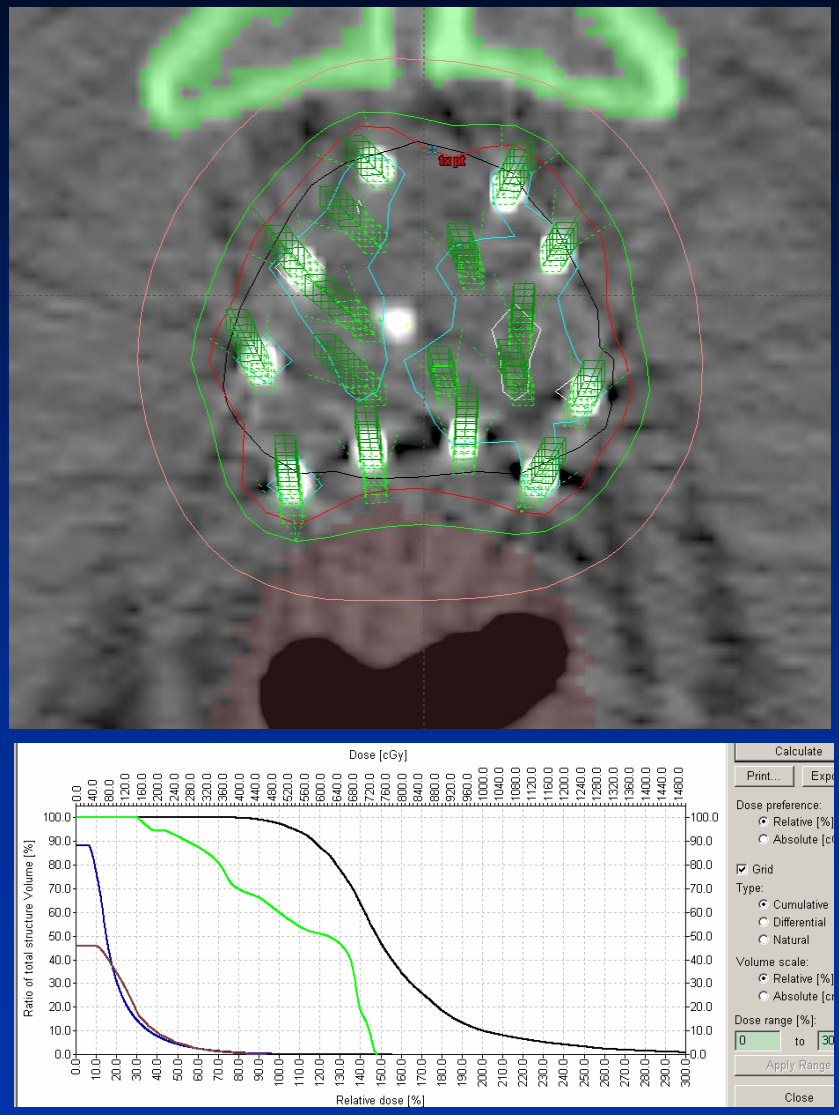




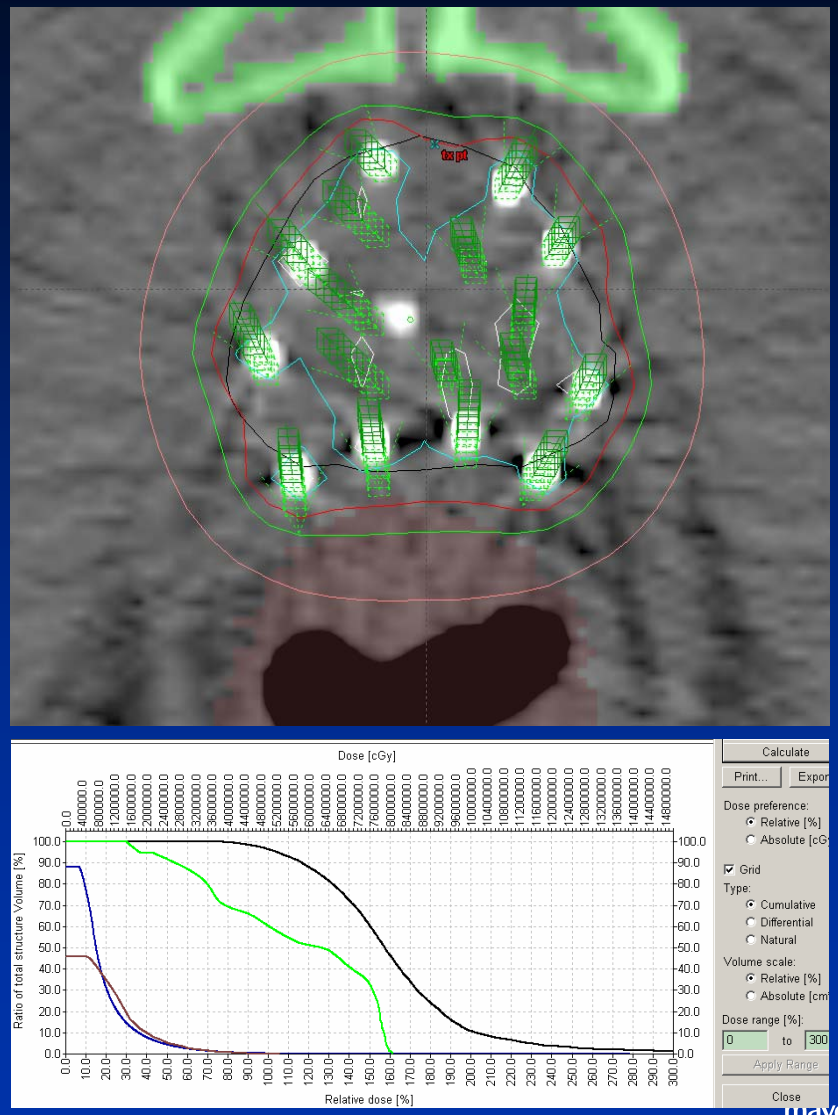


Courtesy Michael Mariscal, Varian

Dose optimization

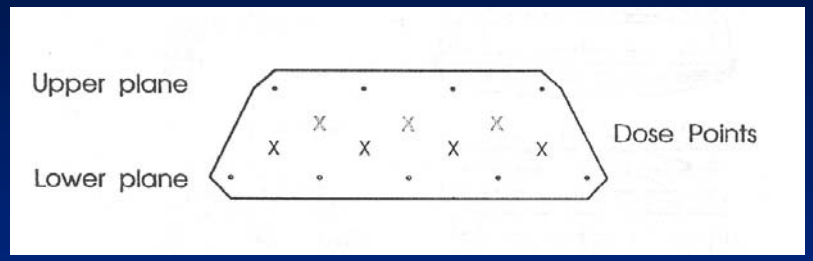


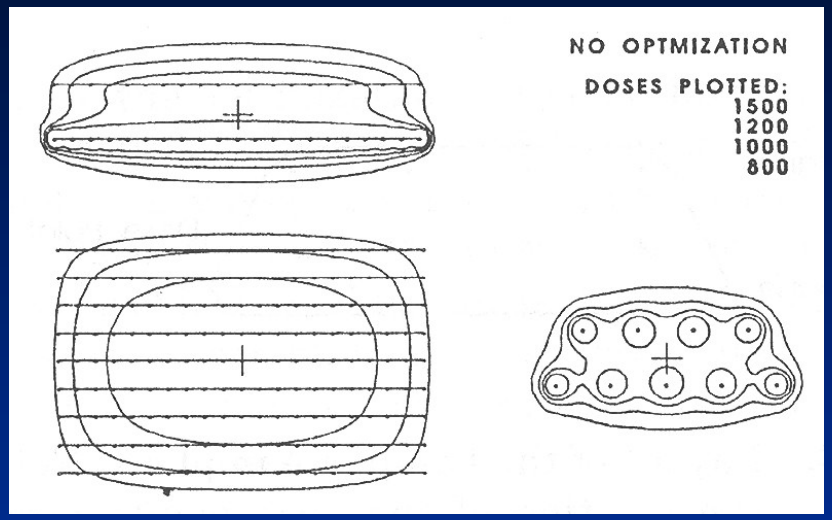
Geometric optimization

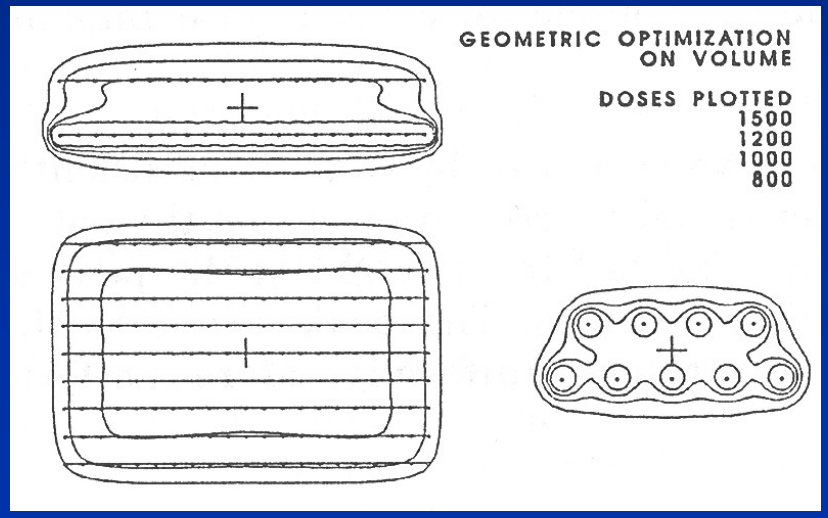


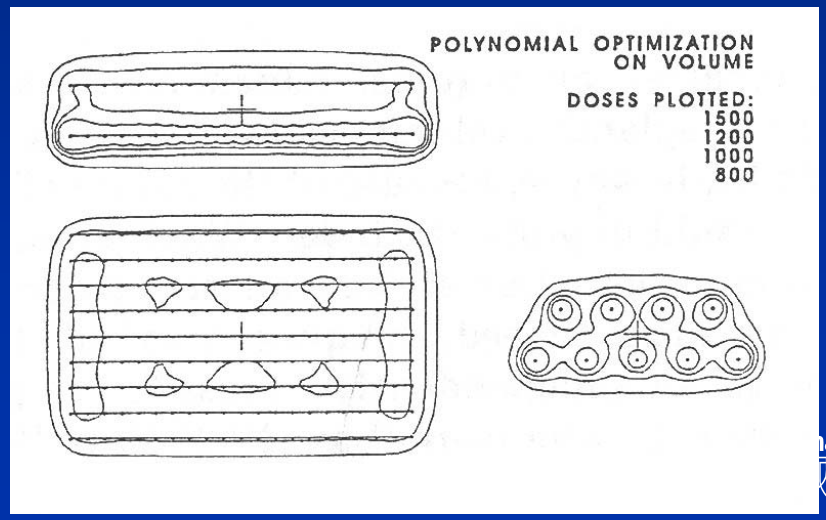
Courtesy Michael Mariscal, Varian

Two-plane breast template (Nucletron)



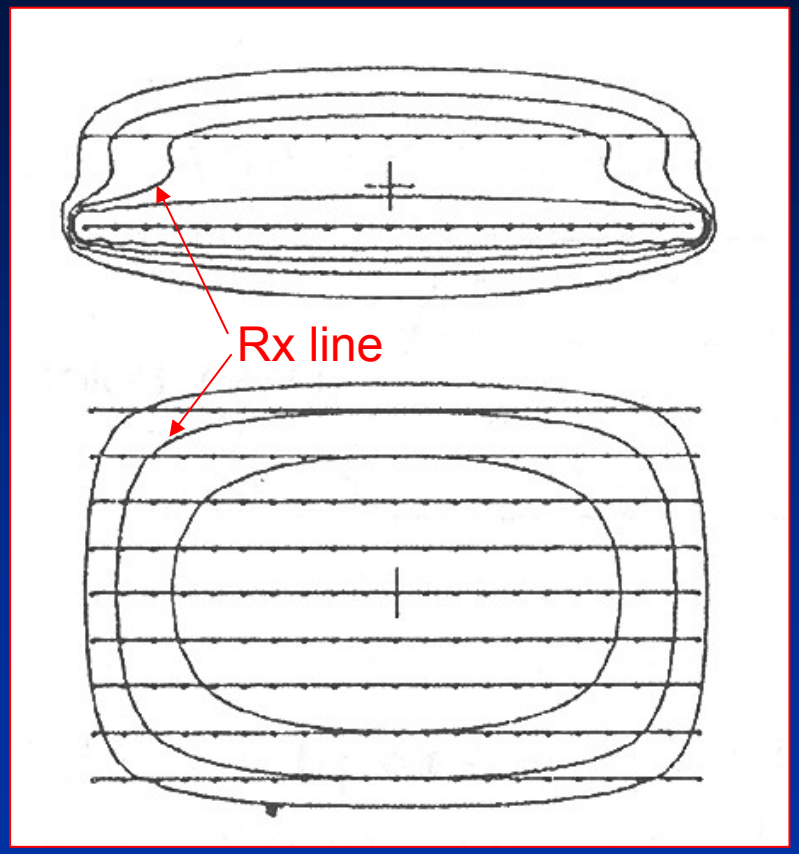




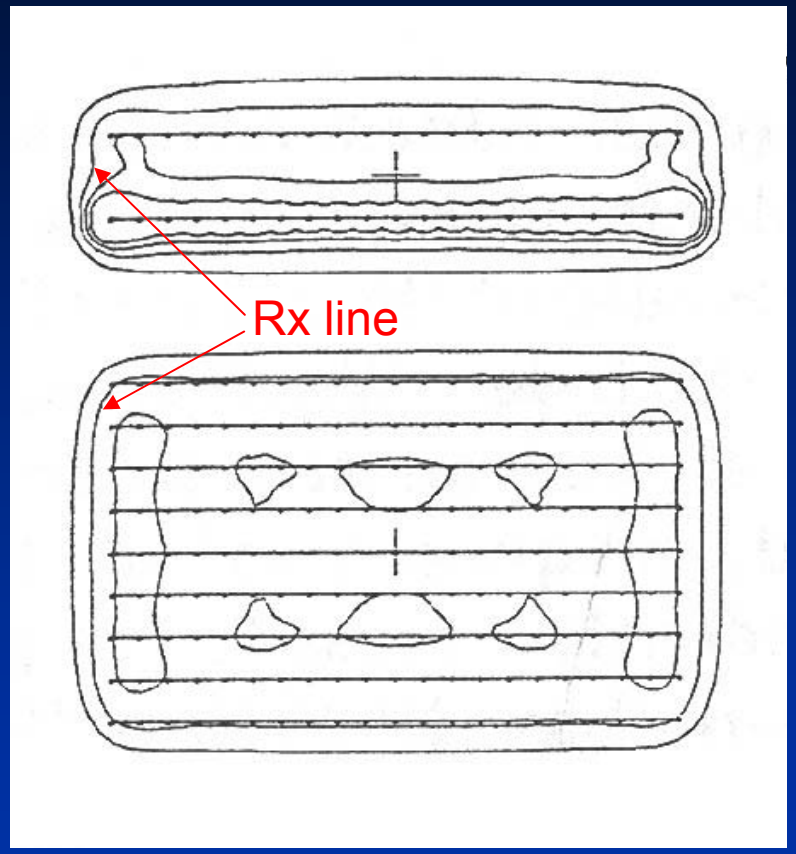


From 1994 brachytherapy summer school

Not optimized



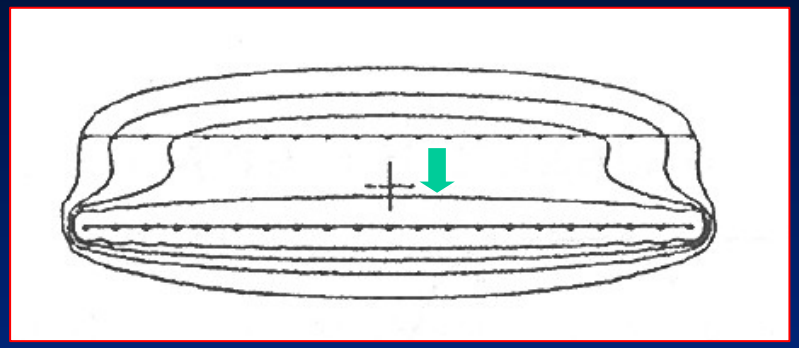
Optimized



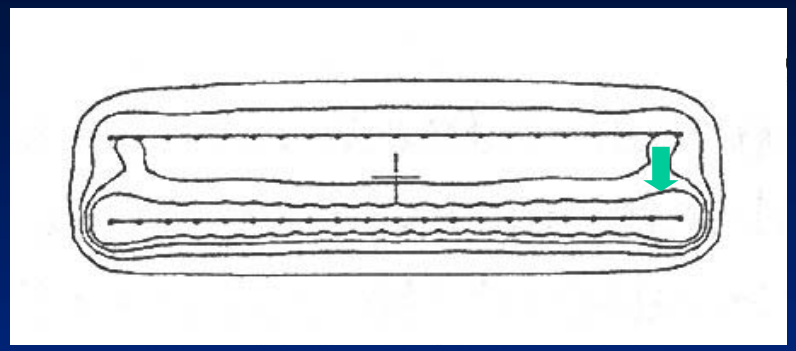
Unoptimized implants must extend <u>beyond</u> target, optimized do not

From 1994 brachytherapy summer school

Not optimized



Optimized



Optimized implants may have locally high doses at the implant boundaries; does this make clinical sense?

"Dose uniformity" in brachytherapy never exists and is its chief advantage; think carefully about what you want; maybe differential dosing is preferable

From 1994 brachytherapy summer school

Brachytherapy generalization

- Inverse square is king.
  - Dwell time optimization is more effective at reducing hot spots than fixing cold spots
  - Plans with more sources are more robust and forgiving of placement errors than plans with fewer

mayo

Summary

- Optimization in brachytherapy always begins with a numerical model of the clinical problem
  - Model of the implant itself and the biological structures it resides in and near
  - Model of what is to be accomplished; this is encoded in the objective function
- The optimizing algorithm searches through potential solutions to find one that best minimizes (or maximizes) the objective function

mayo

- If the modeling has been done well, then that solution will be a good clinical solution, with luck the best available
- However, since the models are always imperfect and approximate, the results need to be evaluated according to criteria both objective and subjective

mayo

- There is seldom a single, best answer
- There usually are multiple objectives that compete with each other
- The planner uses whatever tools there are to steer the optimizer along the Pareto front, developing a set of possible solutions that cannot be improved upon in any one area without losing quality in another
- Our job is to define those choices, and the clinician's is to choose between them

mayo