BMC Urology

Software

A simple computer program for calculating PSA recurrence in prostate cancer patients

Zhongyue Liao[1,2] and Milton W Datta*[2]

Address: [1]Bioinformatics Program Medical College of Wisconsin, Milwaukee, WI, 53226, U.S.A and [2]Department of Pathology Medical College of Wisconsin, Milwaukee, WI, 53226, U.S.A

Email: Zhongyue Liao - zliao@mcw.edu; Milton W Datta* - mdatta@mcw.edu

* Corresponding author

Abstract

Background: Prostate cancer is the most common tumor in men. The most commonly used diagnostic and tumor recurrence marker is Prostate Specific Antigen (PSA). After surgical removal or radiation treatment, PSA levels drop (PSA nadir) and subsequent elevated or increased PSA levels are indicative of recurrent disease (PSA recurrence). For clinical follow-up and local care PSA nadir and recurrence is often hand calculated for patients, which can result in the application of heterogeneous criteria. For large datasets of prostate cancer patients used in clinical studies PSA measurements are used as surrogate measures of disease progression. In these datasets a method to measure PSA recurrence is needed for the subsequent analysis of outcomes data and as such need to be applied in a uniform and reproducible manner. This method needs to be simple and reproducible, and based on known aspects of PSA biology.

Methods: We have created a simple Perl-based algorithm for the calculation of post-treatment PSA outcomes results based on the initial PSA and multiple PSA values obtained after treatment. The algorithm tracks the post-surgical PSA nadir and if present, subsequent PSA recurrence. Times to PSA recurrence or recurrence free intervals are supplied in months.

Results: Use of the algorithm is demonstrated with a sample dataset from prostate cancer patients. The results are compared with hand-annotated PSA recurrence analysis. The strengths and limitations are discussed.

Conclusions: The use of this simple PSA algorithm allows for the standardized analysis of PSA recurrence in large datasets of patients who have undergone treatment for prostate cancer. The script is freely available, and easily modifiable for desired user parameters and improvements.

Background

Prostate cancer is the most common tumor in men and accounts for 41% of male new cancer diagnoses and 14% of cancer deaths. This tumor is predominantly found in older males and thus most studies of prostate cancer progression have been confounded by age-related death not attributable to prostate cancer. For this reason tumor markers for prostate cancer are of utmost importance in estimating death from disease. Prostate cancers arise from the epithelial component of the prostate, a cell that secretes a protein called Prostate Specific Antigen (PSA)[1]. This protein is present within prostatic epithelial cells and is secreted in seminal fluid. In addition, prostatic carcinomas release PSA, which is taken up into the

blood stream. Thus PSA can be measured in the blood of patients and has been used as an effective screening marker for the development of prostate carcinoma in men[2]. Levels of PSA in the blood drop to less than measurable levels (PSA nadir) after surgical removal of the prostate (radical prostatectomy) or treatment of prostate cancer by radiation [3-5]. This reduction of the PSA blood levels can be tracked over time in men, and the finding of increasing levels of serum PSA is considered evidence of a clinical recurrence of prostate cancer (PSA recurrence), thus triggering additional treatment[3]. These values often predate the clinical evidence of prostate cancer recurrence as determined by radiographic or physical examination, and often are the only initial evidence of prostate tumor progression[6]. Thus trends or changes in prostate specific antigen expression over time can be used as a surrogate marker for prostate cancer related morbidity in clinical studies. This is of particular value due to the protracted course of prostate cancer and the current development of chemotherapeutic, radiotheraputic, cryotheraputic, or nutritional interventions designed to delay the course of disease. For these reasons it is very important to be able to accurately predict and identify trends in the serum PSA levels. Yet in most clinical studies and local treatment decisions PSA nadir and recurrence are hand calculated. This often results in a heterogeneous application of PSA nadir and recurrence guidelines, providing a confounding variable for subsequent use of the data in clinical studies. Thus there is a need for processes that allow for the uniform application of PSA nadir and recurrence criteria that can subsequently be used in clinical studies. Here we present a simple Perl script that can be used with patient data for the determination of PSA nadir and recurrence in men who have been treated for prostate cancer. This algorithm orders the serum PSA values within a data set by date, determines the rate of decrease and the post-treatment disappearance of serum PSA (PSA nadir) based on the PSA half-life, and subsequently identifies increases in serum PSA (PSA recurrence). Timelines are calculated from the PSA nadir in months recurrence free or months to PSA recurrence. The clinical standards and practice guidelines used to determine the rules of the algorithm are described, but can be freely and easily modified by the user. The algorithm is demonstrated using a sample prostate cancer dataset.

Methods
The PSA script takes a series of dated PSA values and calculates the PSA drop based on the half-life of PSA in serum to determine PSA nadir. The time to PSA nadir is determined based on the "initial PSA", defined as a PSA value prior to patient treatment, and an estimated PSA half-life of 2–3 days. If PSA nadir is achieved the script subsequently calculates the PSA recurrence, if present, based on sequential rising PSA values and threshold cut-

offs. Requirements for evaluation by the PSA script include the initial dated PSA value prior to treatment, the date of initial treatment, and PSA values after radical prostatectomy with their associated dates of testing. With post-treatment PSA values the script will attempt to determine a PSA nadir (see below) or post the samples as "recurrence status unclear." The PSA script uses the initial PSA value before treatment to calculate the allowed time to PSA nadir (1 month for PSA values less than 50, 3 months for PSA values greater than 50, based on a serum PSA half-life of 2–3 days). If there is an invalid date, defined by a year before 1900, or a month not 1–12, or the follow-up dates are after the current date, then the script will output "Invalid follow-up date found" for that case. If for a particular case there is no (or an invalid) date of treatment, or if there is no initial PSA value then the script will output "unknown". The lack of PSA values within the initial 3 months after treatment will result in the script outputting "nadir unclear" and then using any additional PSA values to attempt a calculation of the PSA recurrence status. The PSA script orders the post-treatment PSA samples by date and takes the PSA values within 1 or 3 months of the treatment date and examines their decrease to zero (PSA nadir). These PSA values must decrease to less than 0.4, but again the user through simple script edits may modify this. Increasing post-treatment PSA values before the PSA nadir value is achieved results in the PSA script output "post prostatectomy elevated PSA." If during the 1 or 3 month post-treatment window the PSA values do not drop below 0.4, then the script outputs "nadir unclear".

Once this nadir value is achieved the date is then stored as the "PSA nadir" where no residual prostate cancer is present. After achieving PSA nadir subroutines are used to identify subsequent PSA increases over time as an indication of PSA recurrence. There are many methods for defining "PSA recurrence", with some authors using any single value above 0.2[7]. Other authors, in particular the American Society for Theraputic Radiology and Oncology (ASTRO), have defined three consecutive increases in serum PSA post-surgery[8]. We have chosen to integrate both systems, with PSA recurrence in the algorithm defined as a single PSA value of greater than 0.4, or a PSA value greater than 0.2 with additional subsequent increasing values. While this has been successful in our hands, the algorithm is designed such that a few simple coding changes can allow a person to alter the algorithm to suit their individual needs. As the PSA levels rise the date of PSA recurrence is determined as the date of the initial PSA rise (either the date of the single value of greater than 0.4 or the date of the PSA value greater than 0.2, before the subsequent rising PSA values) and this date is subtracted from the date of initial PSA nadir to determine the months to PSA recurrence. If there was no documented PSA nadir date then the script uses the initial treatment

date as the nadir date. These time values are subsequently recorded for output. Minimum default values used for the algorithm include a single post-nadir PSA value, otherwise the output of "nadir achieved, recurrence status unknown." is provided. Thus this script can be used as a simple method for calculating PSA recurrence values on any database that tracks serum PSA levels in patients having been treated for prostate cancer.

The script is designed to read data from a flat file in csv format, and can be used by invoking the following file PSARecurReadFile.pl InputDataFileName OutputDataFileName [MaxNumberOfFollowups], where the maximum number of follow-up PSA values per case can be set by the user. If the maximum number of follow-up PSA values is not stated the script will default to a number by calculating the number from the number of values in the input file's title line. For example, if the title line is "init_value, pstm month, pstm year, value A, month A, year A, value B, month B, year B" 2 (value A and value B) will be set as the default maximum number of follow-ups by the script). An output file is provided as additional columns (table 1) with each row representing an individual case. If these columns already exist in the output file, they can be overwritten with new data. This is useful when the script is re-run on a set of existing cases after additional PSA data has been collected.

Table 1: Added output columns for each case

psa_recur_status
psa_recur_time
psa_recur_month
psa_recur_year
psa_recur_free
psa_nadir_month
psa_nadir_year
psa_followup_time

Results

Sample outputs analyzed with the PSA recurrence algorithm demonstrate the calculation of PSA recurrence rates for 30 patients with prostate cancer treated by radical prostatectomy. An anonymous de-identified sample dataset of 30 patients with a range of 0 to 21 post-treatment PSA values per patient were analyzed with the PSA algorithm. Using this dataset (see supplemental file 2) the PSA script correctly calculated 14 cases that underwent PSA nadir. Of these cases 8 were without evidence of PSA recurrence, while 5 underwent PSA recurrence. In 1 case there was insufficient data to calculate PSA recurrence status after achieving PSA nadir. In 10 cases there was insufficient data to calculate PSA nadir due to a lack of PSA values within the 1 or 3 month interval after treatment,

but data was available to calculate PSA recurrence status. The PSA script was not able to calculate the PSA status of 6 cases in the dataset due to missing initial PSA data (3 cases), dates (1 case) of data of any kind (2 cases). The results obtained from the PSA script was in exact agreement with a hand annotated results calculated by one of the authors (M.W.D.). In the cases where the PSA script was not able to provide a definitive result, the author was able to do no better.

Discussion

The use of the PSA script was able to accurately identify patients who had undergone PSA nadir, and patients with subsequent PSA recurrence. In each of these cases the use of a uniform standard was of great value in the subsequent analysis of the outcomes data. A specific area worthy of comment includes the identification of cases with "post-treatment elevated PSA". In these cases there has been a failure to achieve PSA nadir, as defined by decreasing PSA values to less than 0.4 post-prostatectomy. Reasons why a patient may not undergo PSA nadir include the presence of residual prostate cancer within the patient. This may be due to incomplete surgical excision of tumor, spread of the tumor outside the prostate prior to surgery or radiation treatment. In addition, if residual normal prostate tissue is left within the patient after surgery, this may account for residual small but elevated PSA levels. Further evaluation of patients who fail to undergo PSA nadir may identify the potential roles of these factors in the elevated PSA values. After PSA nadir, the elevation of PSA values is indicative of tumor recurrence. The rate at which the PSA rises, or PSA velocity, has been noted to be different between local tumor recurrence and the growth of metastatic disease[9]. This is important as the types of treatment offered to the patients (local radiation or cryotherapy vs. hormonal/chemotherapy) differ. Further modifications of this simple script should focus on improving these measurements if further clinical value is desired.

Conclusions

Here we have presented a simple Perl script for the evaluation of PSA status in patient datasets. Based on the current criteria for PSA nadir and recurrence, the script provides for the uniform application of PSA nadir and recurrence criteria. At the same time the script is flexible enough to allow users to change the criteria (PSA cutoff values, etc) for specific use interests and studies. It is hoped that this will facilitate the use of large patient samples for prostate cancer studies. The Perl script provided without any restrictions and may be modified for any purpose. It is readily available and is attached to this publication (appendix). While the script has been used with an associated Oracle 8 database, the script has been modified

to be used with an input flat file, and as such does not need any database.

Competing interests
None declared.

Author's contributions
The authors (Z.L. and M.W.D.) have contributed equally this work.

Additional material

> ### Additional File 1
> *Additional files are provided and include the Perl PSA Recurrence algorithm along with example files of input data and expected output data for the optimization of the algorithm by individual users.*
> Click here for file
> [http://www.biomedcentral.com/content/supplementary/1471-2490-4-8-S1.csv]

Acknowledgements

References
1. Oesterling JE, Chan DW, Epstein JI, Kimball A. W., Jr., Bruzek DJ, Rock RC, Brendler CB, Walsh PC: **Prostate specific antigen in the preoperative and postoperative evaluation of localized prostatic cancer treated with radical prostatectomy.** *J Urol* 1988, **139:**766-772.
2. Chu TM, Murphy GP: **What's new in tumor markers for prostate cancer?** *Urology* 1986, **27:**487-491.
3. Partin AW, Pound CR, Clemens JQ, Epstein JI, Walsh PC: **Serum PSA after anatomic radical prostatectomy. The Johns Hopkins experience after 10 years.** *Urol Clin North Am* 1993, **20:**713-725.
4. Chodak GW, Neumann J, Blix G, Sutton H, Farah R: **Effect of external beam radiation therapy on serum prostate-specific antigen.** *Urology* 1990, **35:**288-294.
5. Cadeddu JA, Pearson JD, Partin AW, Epstein JI, Carter HB: **Relationship between changes in prostate-specific antigen and prognosis of prostate cancer.** *Urology* 1993, **42:**383-389.
6. Moul JW: **Prostate specific antigen only progression of prostate cancer.** *J Urol* 2000, **163:**1632-1642.
7. Polascik TJ, Oesterling JE, Partin AW: **Prostate specific antigen: a decade of discovery--what we have learned and where we are going.** *J Urol* 1999, **162:**293-306.
8. Shipley WU, Thames HD, Sandler HM, Hanks GE, Zietman AL, Perez CA, Kuban DA, Hancock SL, Smith CD: **Radiation therapy for clinically localized prostate cancer: a multi-institutional pooled analysis.** *Jama* 1999, **281:**1598-1604.
9. Partin AW, Pearson JD, Landis PK, Carter HB, Pound CR, Clemens JQ, Epstein JI, Walsh PC: **Evaluation of serum prostate-specific antigen velocity after radical prostatectomy to distinguish local recurrence from distant metastases.** *Urology* 1994, **43:**649-659.

Pre-publication history
The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2490/4/8/prepub