

INFOPOINTS

A web-based system for individualised survival estimation in breast cancer

Clinicians want prognostic tools that not only aid prognostic classification, but also give quantitative probabilities of survival.¹ We describe a way of generating survival estimates that uses existing survival data and generates survival curves online dynamically.

The website

On the website <http://finprog.primed.info> a selection of prognostic factors are available for case-match survival estimation (figure).^{*} The default selection in the drop down list for each factor is "all," which means that no selection has been made for the specific factor. The user can enter a prognostic factor profile by selecting any of the categories in the drop down lists. The software then queries the database to retrieve data on patients with matching prognostic profiles and known outcome and calculates a survival curve according to the Kaplan-Meier product-limit method using the actual survival data of all matching patients. The number of patients at risk, the confidence intervals for the Kaplan-Meier estimates, and the median survival time are also displayed. The user can compare two factor profiles by clicking the "two profiles" option. The distribution of patients according to vital status, therapy received, or a specific prognostic factor can also be displayed as a table or a chart (figure). The website also contains basic information about survival statistics and the prognostic factors, including guidelines for selecting variables and interpreting the results.

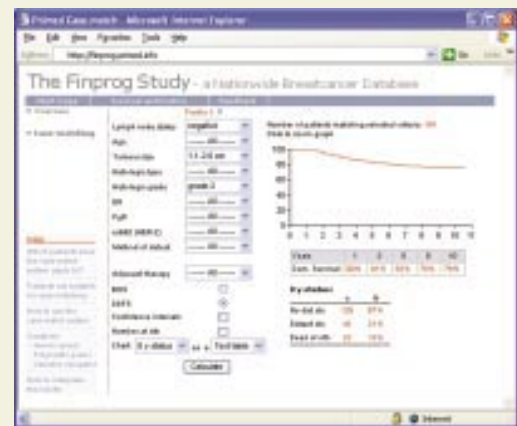
The database

This case-match survival estimation system could be applied to any clinical database with time to event information. At our website it is applied to the FinProg breast cancer series,² which includes individual clinical data on women diagnosed with breast cancer in 1991-2 in five regions of Finland (representing about 50% of the Finnish population). The minimum prognostic factor information was available for 2842 patients (91% of all cases in the five regions and 51% of all breast cancers diagnosed in Finland in 1991-2). After we excluded some cases (in situ carcinoma, bilateral tumour, etc), 2032 patients were left in the final data set. All personal identification information was deleted before linking the database to the website. The distant disease free survival is calculated from the date of diagnosis to the occurrence of metastases outside the locoregional area and survival to death from breast cancer. Patients who died of other causes or were lost to follow up are censored.

The prognostic factors available on the website are those recommended by the National Institutes of Health Consensus Development Panel³ and the International Consensus Panel on the Treatment of Primary Breast Cancer.⁴

The advantages

The advantage of this system is that clinicians and researchers can obtain survival estimates based on actual data, rather than inferential estimates generated by a regression formula. The output is a survival curve for the entire available follow up period and not just for a single time point. However, the robustness and accuracy of the Kaplan-Meier estimates depend on the quality of the underlying data set.



Users are not restricted to a single model but can enter any prognostic factor data they have available. However, the more variables selected the fewer patients will match the selected categories and the more uncertain the survival estimates. Users should therefore study the variable selection guide on the website and select the most important variables first, keeping an eye on the number of matching patients and the confidence intervals.

Our system also responds to a demand for openness, giving other researchers the possibility to check not only published results but also to further explore the database. It offers a compromise between demands to make raw data available and the reluctance of many researchers entirely to release their valuable datasets. This interactive, web based system may facilitate explorative analyses of prognostic factors and could be applied to a variety of diseases.

Johan Lundin *researcher*

Mikael Lundin *researcher*

Department of Oncology, University of Helsinki, FIN-00290, Helsinki, Finland

Jorma Isola *professor*

Institute of Medical Technology, University of Tampere and Tampere University Hospital, FIN-33014 Tampere, Finland

Heikki Joensuu *professor*

Molecular/Cancer Biology Research Programme and Department of Oncology, University of Helsinki, FIN-00290, Helsinki, Finland

The breast cancer prognostic system is available at <http://finprog.primed.info>.

1 Kattan M. Statistical prediction models, artificial neural networks, and the sophism "I Am a Patient, Not a Statistic." *J Clin Oncol* 2002;20:885-7.

2 Lundin J, Lundin M, Holli K, Kataja V, Elomaa L, Pylkkanen L, et al. Omission of histologic grading from clinical decision making may result in overuse of adjuvant therapies in breast cancer: results from a nationwide study. *J Clin Oncol*;19:28-36.

3 National Institutes of Health Consensus Development Conference statement: adjuvant therapy for breast cancer, 1-3 November, 2000. *J Natl Cancer Inst Monogr* 2000;30:5-15.

4 Goldhirsch A, Glick JH, Gelber RD, Coates AS, Senn HJ. Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. Seventh International Conference on Adjuvant Therapy of Primary Breast Cancer. *J Clin Oncol* 2001;19:3817-27.

