Utility of prognostic scoring systems for colorectal liver metastases in an era of advanced multimodal therapy

E. Gregoire, E. Hoti, D.L. Gorden, G. Pascal, D. Azoulay

To cite this version:

HAL Id: hal-00599230
https://hal.archives-ouvertes.fr/hal-00599230
Submitted on 9 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
UTILITY OF PROGNOSTIC SCORING SYSTEMS FOR
COLORECTAL LIVER METASTASES IN AN ERA OF ADVANCED
MULTIMODAL THERAPY

E Gregoire MD\textsuperscript{1}, E Hoti MD, FRCS\textsuperscript{1,2}, DL Gorden MD\textsuperscript{1,3}, G Pascal MD\textsuperscript{1}; D Azoulay MD, PhD\textsuperscript{1}

1) AP-HP Hôpital Paul Brousse, Centre Hépato-Biliaire, F-94804 Villejuif, France
2) Liver Transplant Unit, Saint Vincent University Hospital, Elm Park, Dublin 4, Ireland
3) Division of Hepatobiliary Surgery and Liver Transplantation, Vanderbilt University Medical Center, Nashville, TN, USA

Corresponding author / reprint requests:
Professor Daniel Azoulay
AP-HP Hôpital Paul Brousse, Centre Hépato-Biliaire
12 Avenue Paul Vaillant Couturier, F-94804 Villejuif, France
Tel: 00 33 145593036
Fax: 00 33 145593857
E-mail: daniel.azoulay@pbr.aphp.fr

Key words: Colorectal liver metastases
Prognostic score
Validity confirmation
ABSTRACT

OBJECTIVES: To assess the general applicability of prognostic scores for colorectal liver metastases (CRLM).


RESULTS: Six scoring systems and fourteen prognostic factors within these studies were identified. No prognostic factor was common in all scoring methods. Five scores retained the number of metastases as a prognostic factor. Size of metastases and time between the onset of the primary tumor and the discovery of metastases were present in four scores. Three scores predicted 5-year survival using carcinoembryonic antigen (CEA) and R1 resection. Only two scores were assessed preoperatively. Successive scoring methods had improved predictive accuracy compared to earlier systems. However, their applicability in general populations remains debatable. An evaluation of the scores applicability to different patient populations demonstrated that the models were minimally effective in predicting disease-specific survival and recurrence, suggesting that stratification of patients by clinical and pathologic factors alone, may be clinically unreliable and not applicable for selection of patients for surgery.

CONCLUSION: The utility of prognostic models on general populations is inconsistent. Current clinicopathologic factors may be inadequate to determine disease prognosis in CRLM. Future attempts to develop prognostic scores should include additional biologic and clinical variables, and be validated in larger populations.

Abstract word count: 221
INTRODUCTION

Colorectal cancer is the second most common cause of cancer related deaths in Europe (1) and the fourth most common cause in United States (2). Approximately, 1.2 million cases of colorectal cancer occur yearly worldwide, with 412,900 new cases diagnosed in Western Europe alone (14.4% of all newly diagnosed cancer cases) (2). Between 15 to 25% of patients have hepatic metastases at the time of their initial presentation (3, 4, 5). In addition, 40 to 50% of patients with initially localized disease will eventually develop metachronous hepatic metastases after primary tumor resection (3). Only 20% of these patients will be immediately amenable to surgery (6). Hepatic resection remains the sole potentially curative treatment, able to achieve long-term survival with a 5-year survival ranging from 37 to 58% (7). However, even in patients who undergo liver resection with curative intent, 60% will develop recurrent hepatic metastases. Of these, one third will benefit from repeat hepatectomies with reported 5-year survival rates ranging from 26 to 41% (8). Modern chemotherapy can offer improved short and long term survivals, but 5-year survivors without concomitant surgical resection are the exception (9).

Several groups have identified and used various prognostic criteria to ensure a maximum treatment benefit for patients selected for hepatic resection. Prognostic scoring systems have been employed as early as the 1980s with the aim to better select candidates for surgery, stratify patients for clinical trials, predict tumor recurrence risk and lastly for postoperative surveillance. With
increasingly effective systemic chemotherapies, both cytotoxic and targeted as well as advances in surgical techniques, that have significantly impacted patients survival with colorectal cancer metastases, the utility of currently available prognostic scoring systems is important to assess.
METHODS

Search strategy
We conducted a systematic literature review to assess the accuracy of prognostic scoring systems used to identify good candidates for hepatic surgery or groups with favorable prognostic factors.

A computerized search was made using Medline and Embase database from January 1980 to 2008. We selected English language articles and human studies as the search limiting criteria. The Mesh heading “colorectal neoplasms” in combination with keywords “liver metastases” and “liver resection” yielded 306 hits in Medline and 2499 in Embase. The keyword « prognostic scoring system » led to 3024 hits in Embase and 759 in Medline, whereas « prognostic score » yielded 5811 hits in Embase and 5472 in Medline. The combination of those keywords allowed us to select 1859 hits in Embase and 354 in Medline. We organized the abstract selection from the databases using a repetitive search strategy. Combinations of key word search terms included « prognostic score » and « colorectal neoplasms liver metastasis liver resection » then « prognostic scoring system » and « colorectal neoplasms liver metastasis liver resection » and finally « prognostic score prognostic scoring system » and « colorectal neoplasms liver metastasis liver resection ». These 3 search orientations produced 58 hits in Embase and 18 in Pubmed.
We excluded duplicate articles and obtained 30 abstracts. These abstracts were then reviewed in detail. Letters, reviews without original data, studies with less than 200 patients were in turn excluded. Twenty one articles were rejected upon examination of the abstract. The exclusion process yielded 9 articles. Manual searching of the reference lists of these articles identified 3 additional relevant studies, yielding a total of 12 studies. Evaluation of the complete manuscripts led to further elimination based on heterogeneous populations of non-surgically treated patients, number of patients below 200, non relevant data. These exclusions produced a final study group of 6 non-duplicated series.
RESULTS

The principal prognostic scoring systems

Only 6 of the 30 articles met all eligibility criteria for the review. The studies included in the analysis are shown in Table 1.

Nordlinger et al (10) in 1996 reported a collected series (multicenter) of 1568 patients with resected colorectal liver metastases. Seven parameters were found to be independent risk factors including: age, size of the largest metastasis and CEA level, stage of the primary tumor, disease free interval from primary to metastasis (DFI), number of liver nodules and resection margin. One point was allocated to each prognostic factor and the prognostic scoring system divided the population into 3 risk groups with 2-year survival rates ranging from 79% (0-2 points) to 43% (5-7 points). The French group advised to take into account the operative risk and the risk group when considering hepatic resection.

Fong et al (11) in 1999 reported a retrospective study involving 1001 patients. Seven factors were found to be independent predictors of poor prognosis: positive margin, extra-hepatic disease, node positive primary, disease free interval, number of liver nodules, size of the largest metastasis and CEA level. One point was allocated to each risk factor. No patient with a score of 5 was a long-term survivor. According to Fong et al, patients with up to two risk factors can have a favourable outcome after hepatic resection whereas patients with 3 or more factors should be considered for systemic therapy and clinical trials.
Iwatsuki et al (12) in 1999 developed the Pittsburgh risk score. The independent negative prognostic indicators were: positive surgical tumor involvement including lymph nodes, margins, extra hepatic tumor number of three or more, bilobar tumors and time from primary tumor treatment to the diagnosis of hepatic metastases of 30 month or less. After exclusion of patients with positive margins or extra hepatic tumor, the remaining significant criteria were: tumor number of three or more, tumor size greater than 8 cm, time from primary tumor treatment to hepatic metastasis assessment of 30 month or less and bilobar tumors. Five prognostic groups were identified ranging from group 1, without any risk factor (48% 5-year survival), to group 5 with 4 risk factors (0% 5-year survival). Group 6, which consisted of patients with positive margins or extra hepatic tumor was not considered for liver resection.

More recently, Zakaria et al (13) identified tumor size, node positive primary, DFI, and positive hepatoduodenal nodes as major determinants of survival and recurrence. Three risk groups were described with 5-year survival between 20 and 55%.

Malik et al (14) in 2007 reported a preoperative prognostic score including only the number of metastases and the presence or the absence of an inflammatory response to tumor. Both factors were found to influence both overall and disease-free specific survival on multivariate analysis. Three risk groups derived. The 5-year survival of patients scoring 0 was 49% compared to 34% for patients scoring 1. None of the patients that scored 2 were alive at 5 years. This
score allows clinicians to determine before surgery which patient will have a favourable prognosis as well as identify patients in whom surgical resection may offer little additional benefits compared to palliative chemotherapy.

Rees et al (15) in 2008 reported a single centre series of 929 patients. Seven independent risk factors were found to be predictors of poor survival: number of hepatic metastases > 3, node positive primary, poorly differentiated primary, extra hepatic tumor, tumor diameter > 5cm, CEA level > 60ng/mL and positive surgical margins. Patients with the worst prognostic postoperative criteria had a 5-year cancer specific survival of 2% and those with the best prognostic postoperative criteria had a 5-year survival of 64%. The authors concluded that the Basingstoke predictive index could be used for risk-stratifying patients who may benefit from intensive surveillance or to select adjuvant therapies or trials.
Common prognostic criteria

All together, 14 risk factors were identified among the different scoring systems. The main result of the present analysis is that no prognostic factor was common for all the prognostic scores and no scoring system regrouped all the prognostic factors.

All the authors, except Zakaria et al (13), regarded the number of metastases as prognostic factor, however, different cut-offs were used (from one [11] to more than eight metastases [14]).

Only four of the six scores considered the disease free interval between primary tumor and the onset of liver metastases as prognostic factor (10, 11, 12, 13). Size of the metastases was considered as a prognostic factor in five scores (10, 11, 12, 13, 15). Finally, three of the scores predicted survival based on CEA level, positive surgical margins and node positive primary (10, 11, 15).

The size of the metastases and CEA level were variably expressed in prognostic scores: size ranging from 5 (10, 15) to 8 cm (12, 13, 14) and CEA level from 30 to 200 ng/mL (10, 11, 15). Only Nordlinger et al (10) identified age above 60 years as a prognostic factor in multivariate analysis. Extra hepatic disease was a contraindication to hepatic surgery for Fong et al (11) and Iwatsuki et al (12). Only two scores were available before liver surgery (12, 14). Four scoring systems required postoperative data such as surgical margin or blood transfusion in order to predict survival (10, 11, 13, 15).
DISCUSSION

Metastatic colorectal cancer remains an important health problem in developed countries (Europe and United States) associated with a high human toll as well as substantial financial costs (1, 2).

A number of prognostic criteria have been developed to ensure a maximal benefit in terms of overall and disease free survival for patients selected for hepatic resection (Figure 1). They have been used to predict recurrence risk and patients’ survival, to help clinicians to select patients for resection and/or adjuvant therapy, to stratify patients into risk categories for clinical management and also for postoperative surveillance.

In spite of the significant literature on the subject, selecting the appropriate candidates for hepatic resection for colorectal metastasis remains controversial. Since 1980, six scoring systems, based on large populations, have been described (10, 11, 12, 13, 14, 15). There are significant differences between them, and the general wider applicability of any of the scoring systems has not been clearly evaluated.

To some extent, scores require internal and external validation before general use. This should establish whether the systems/model is applicable to a wider patient population. This has not been consistently done. Malik et al (14), identified a score with a correlation coefficient and a gradient of slope higher than those of the model described by Fong et al. Similarly Iwatsuki et al (12), proposed the Pittsburg scoring system which is more discriminating than the
score described by Nordlinger et al (10), for patients with a score greater than 1. Nevertheless, no consensus exists and one can wonder whether these scoring systems can be really applicable in the general population. To this end, Zakaria et al evaluated whether the proposed risk scoring systems were applicable for the population of the Mayo Clinic (13). In their updated experience with hepatic resections for metastatic colorectal cancer, they imported their findings into 3 scoring systems (Nordlinger et al [10], Fong et al [11], and Iwatsuki et al [12]) in order to determine whether outcomes for risk stratification were similar. Their study showed that the 3 models “were only marginally better than chance alone in predicting disease-specific survival and recurrence”. These observations suggest that stratification of patients by clinical and pathologic factors alone, although statistically sound, may be clinically unreliable and are not widely applicable for selection of patients for surgery or for comparing cohorts of patients between institutions.

There are several widely used statistical approaches to assess the prediction accuracy of a prognostic scoring system.

The internal approach is a procedure restricted to a single data set. One common way to establish whether a model might be highly capable in predicting survival is data splitting or cross-validation. For example, leave-one out cross-validations is a mean to evaluate prediction accuracy. The population will be divided into k subsets of approximately equal size. Than, the total data will be trained k times, each time leaving out one of the subsets from training, but using only the
omitted subset to compute whatever error relevant criterion. If k equals the sample size, this is called “leave one out cross validation”. Bootstrapping is another way. It consists in analysing subsamples (and not subsets) of the data obtained by a random sample with replacement from the full sample (50 to 2000 subsamples might be used). But those validation methods involve bias (16) and in some cases it could be excessively optimistic (17, 18) having cross-validation methods still as an internal procedure. In our review, two of the six scores used bootstrapping validation (10, 15).

The temporal validation consists in an evaluation on a second data set from the same center (19). It is a prospective evaluation, independent of the original data and model-fitting process. However, it can be problematic if the outcome is the survival time. It could need several years to accrue an adequate number of events in a further cohort.

The external validation method is an evaluation on data from one or more other centers. Neither internal nor temporal evaluation addresses the wider issues of the generalization of the model. External validation can be based on retrospective data and so is viable for validating survival models needing long follow-up.

Current scoring systems rely on radiographic and other variables as described. Multimodal treatment however has evolved to include the complementary markers to guide molecular targeted therapy (K-Ras mutation status), novel cytotoxic agents and complex surgical strategies. Medical surgical approaches
that involve treatment of isolated hepatic metastases as well as extrahepatic disease have been shown to be effective. New promising prognostic factors have been recently identified that may improve the predictive power of prognostic scoring systems. For example, the markers of systemic inflammation response described by Leitch et al (20), were associated with poor outcomes. There are also new tools to identify biologic, histological, genetic or proteomic “signatures” as prognostic factors such as micro-array analysis and high throughput screening techniques (21, 22, 23, 24). These new tools have been demonstrated to have high predictive accuracy compared the traditional clinical prognostic factors, hence the biologic characteristics of the tumor and the “host” can no longer be ignored and should be included in future prognostic scores.

Hospital and surgeon volumes should also be considered in prognostic studies. It has been already established that both surgeon-specific procedures and hospital volume predict outcomes for liver resection (25, 26). Obviously, all these finding must be adjusted to hospital mortality and morbidity rates for hepatic surgery in order to correctly evaluate benefits and risks for each patient.

The biology of the tumour is another factor, which cannot be ignored and should be considered in the prognostic studies. For instance, the study by Adam et al. (27) showed that the prognosis of patients demonstrating tumor progression while on chemotherapy was worse when compared to that of patients who responded objectively to chemotherapy. The response to adjuvant chemotherapy
plays a key role in survival outcomes. Measurement of the disease evolution for particular patients will add a dynamic aspect to assessment of prognosis.

Surgical procedures have considerably improved during the past 20 years including portal embolisation (28), two-stage hepatic surgery (29) and extended hepatectomy with hypothermic perfusion (30). New chemotherapy regimens have also allowed for improved resectability of hepatic metastases. In this study, 6 scoring systems derived from retrospective studies ranging from 1968 to 2006. The population of the patients eligible for hepatic resection has changed and those retrospective studies do not consider the influence of these recent advances (Figure 2). Conducting prospective multicentric studies, with long follow-up, is the best way to achieve a consensus. The European registry of hepatic metastases of colorectal origin (LiverMetSurvey) which prospectively collects significant international data concerning the history, treatment (chemotherapy, surgery, combined ablation) and the outcome of operated patients was implemented to this end. As the final purpose of this registry is to analyse prospectively and on a large scale the results of surgically treated patients to define guidelines for an optimal treatment strategy, this will be particularly useful in the development of newer prognostic scores.

In conclusion, at this time, the prognostic scoring systems for selection of patients with colorectal hepatic metastases for resection are inconsistent and conflicting. Given advances in molecular diagnostics and therapeutics as well as
newer chemotherapies, the currently used conventional clinicopathologic factors may be inadequate for the purposes of prognostication.

Colorectal cancer is regarded as a heterogenous disease resulting from disregulation of multiple pathways disease. There has been increased interest in identifying biologic indicators that may help better define patients likely to have a poor clinical outcome (increased disease recurrence, decreased survival). Recent studies have shown that protein markers may be better predictors of outcome in patients with CRC metastases. Parallel to this, the concept of resectable and non-resectable disease has changed completely. The definition of resectability has evolved and yet is not uniform. Currently, the definition of resectability with curative intent is the ability of the surgeon to clear, with a macroscopically negative margin, all detectable malignant disease, while leaving an adequate volume of healthy future remnant liver.

It is necessary that future attempts to develop newer prognostic scoring systems be conducted in larger patient populations and likely will include additional biologic indicators as well as clinical factors. External confirmation for validity prior to being introduced for uses other than counseling patients is also required for any newly developed scoring system.
REFERENCES


<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>N</th>
<th>5 - year survival</th>
<th>Median follow-up</th>
<th>Prognostic factors for disease specific survival</th>
<th>Prognostic score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rees et al (22)</td>
<td>2008</td>
<td>929</td>
<td>36%</td>
<td>26.4</td>
<td>For disease specific survival Number of metastases &gt; 3 Node positive primary Poorly differentiated primary Extra-hepatic disease Tumor diameter ≥ 5cm CEA &gt; 60ng/mL Positive resection margin</td>
<td>The Basingstoke predictive index Post-Operative score 30: 5 year survival 2% Post-Operative score 0: 5 year survival 64%</td>
</tr>
<tr>
<td>Malik et al (21)</td>
<td>2007</td>
<td>687</td>
<td>45%</td>
<td>34</td>
<td>For disease specific survival Number of metastases &lt; or &gt; 8 Inflammatory response to tumor Correlation coefficient and the gradient of slope is higher than the Fongs' score</td>
<td>Preoperative prognostic score 0= ≤ 8 meta and absence of IRT 5 year DS survival: 49% 1= ≥ 8 meta or IRT 5 year survival: 34% 2= ≥ 8 meta and IRT 5 year survival: 0%</td>
</tr>
<tr>
<td>Zakaria et al (19)</td>
<td>2007</td>
<td>662</td>
<td>37%</td>
<td>36</td>
<td>For disease specific survival Blood transfusion Positive hepatoduodenal nodes Disease-free interval from primary &lt; 30 month Metastases size &gt; 8cm</td>
<td>Neither survival nor recurrence was stratified discretely by any of the scoring systems (Fong et al., Iwatsuki et al., Nordlinger et al.) 1= blood transfusion and/or positive hepatoduodenal nodes 5 year DS survival: 55% 2= any risk factor except blood transfusion and/or positive hepatoduodenal nodes 5 year DS survival: 39% 3= positive hepatoduodenal nodes and any risk factor 5 year DS survival: 20%</td>
</tr>
<tr>
<td>Fong et al (17)</td>
<td>1999</td>
<td>1001</td>
<td>37%</td>
<td>32</td>
<td>Positive margin Extra-hepatic disease Node positive primary Disease-free interval from primary &lt; 12 months Number of metastases &gt; 1 Largest metastasis size &gt; 5cm CEA &gt; 200 ng/mL Correlation coefficient and the gradient of slope higher than the number of tumors</td>
<td>Positive margin and extra-hepatic disease were considered as contraindications to liver resection Fong et al Each risk factor is one point 0= 5 year DS survival: 60% 1= 5 year DS survival: 44% 2= 5 year DS survival: 40% 3= 5 year DS survival: 20% 4= 5 year DS survival: 25% 5= 5 year DS survival: 14% (no survivors at 5 years)</td>
</tr>
<tr>
<td>Iwatsuki et al (18)</td>
<td>1999</td>
<td>243</td>
<td>32.3%</td>
<td>32</td>
<td>Number of metastases ≥ 3 Bilobar tumor Disease-free interval from primary ≤ 30 month Tumor diameter ≥ 8cm</td>
<td>Patients with positive margin and extra-hepatic disease (including lymph nodes) excluded Pittsburgh system 1= no risk factor</td>
</tr>
</tbody>
</table>
Nordlinger et al (16) 1996 1568 28% 19

<table>
<thead>
<tr>
<th>Age &gt; 60</th>
<th>Tumor extension into serosa</th>
<th>Node positive primary</th>
<th>Disease-free interval from primary &lt; 24 month</th>
<th>Number of metastases ≥ 4</th>
<th>Largest metastasis size ≥ 5 cm</th>
<th>Margin &lt; 1 cm</th>
<th>1= &lt; 3 risk factors</th>
<th>2 year DS survival: 79%</th>
<th>2= 3 to 4 risk factors</th>
<th>2 year DS survival: 60%</th>
<th>3= 5 to 7 risk factors</th>
<th>2 year DS survival: 43%</th>
</tr>
</thead>
</table>

5 year DS survival: 48,3%
2= 1 risk factor
5 year DS survival: 36,6%
3= 2 risk factors
5 year DS survival: 19,9%
4= 3 risk factors
5 year DS survival: 11,9%
5= 4 risk factors
5 year DS survival: 0%
6= Patients with positive margin and extra-hepatic disease
5 year DS survival: 0%
### Table 1: List of the main prognostic scoring systems (1980 – 2008)

DS survival – Disease specific survival  
IRT – Inflammatory response to tumor  
CEA – Carcinoembryonic antigen
Figure 1: The prognostic scoring systems and the evolution of oncosurgical strategies