Review

Prognostic models for outcome following liver resection for colorectal cancer metastases: A systematic review

L. Spelt a,c, B. Andersson a,c, J. Nilsson a,b,c, R. Andersson a,c,*

* Department of Surgery, Clinical Sciences Lund, Lund University, Lund, Sweden
b Department of Cardiothoracic Surgery, Clinical Sciences Lund, Lund University, Lund, Sweden
c Skåne University Hospital Lund, Lund, Sweden

Accepted 24 October 2011
Available online 12 November 2011

Abstract

Background: Liver resection provides the best chance for cure in colorectal cancer (CRC) liver metastases. A variety of factors that might influence survival and recurrence have been identified. Predictive models can help in risk stratification, to determine multidisciplinary treatment and follow-up for individual patients.

Aims: To systematically review available prognostic models described for outcome following resection of CRC liver metastases and to assess their differences and applicability.

Methods: The Pubmed, Embase and Cochrane Library databases were searched for articles proposing a prognostic model or risk stratification system for resection of CRC liver metastases. Search terms included ‘colorectal’, ‘liver’, ‘metastasis’, ‘resection’, ‘prognosis’ and ‘prediction’. The articles were systematically reviewed.

Results: Fifteen prognostic systems were identified, published between 1996 and 2009. The median study population was 305 patients and the median follow-up was 32 months. All studies used Cox proportional hazards for multi-variable analysis. No prognostic factor was common in all models, though there was a tendency towards the number of metastases, CRC spread to lymph nodes, maximum size of metastases, preoperative CEA level and extrahepatic spread as representing independent risk factors. Seven models assigned more weight to selected factors considered of higher predictive value.

Conclusion: The existing predictive models are diverse and their prognostic factors are often not weighed according to their impact. For the development of future predictive models, the complex relations within datasets and differences in relevance of individual factors should be taken into account, for example by using artificial neural networks.

© 2011 Elsevier Ltd. All rights reserved.

Keywords: Liver resection; Colorectal cancer; Liver metastases; Predictive models; Prognosis; Outcome

Introduction

Yearly, about 1.2 million cases of colorectal cancer (CRC) occur worldwide, with approximately 195 200 newly diagnosed cases in North-Western Europe. It is one of the most common malignancies and the third leading cause of cancer-related mortality in both men and women. The general 5-year survival rate is estimated to be 65%, declining to 11% if distant metastases are present.

Within three years after the diagnosis of CRC, 29% of the patients will develop liver metastases. This condition will be fatal if left untreated, with a median survival measured in months. For patients with resectable liver metastases, a resection with curative intent is the treatment of choice, which renders a reported 5-year survival rate in the range of 38–58%.

By means of novel treatment strategies, such as portal vein embolisation (PVE), two-stage hepatectomy...
and preoperative down-sizing chemotherapy, together with more “aggressive” surgery due to increased safety of the procedure per se, the number of patients who are suitable for resection has gradually increased.11–15

Several previous study groups have examined prognostic factors for tumour recurrence and survival after liver resection for CRC metastases.16–20 The development of predictive models helps to stratify patients into risk categories for the selection of management strategies, to predict prognosis and possibly also to allow a more accurate evaluation of the influence and effectiveness of new therapeutic interventions. Currently, few such models are available, of which the most well known ones are those by Fong et al.,21 Rees et al.8 and Nordlinger et al.22 However, as is the case with these three, most models do not take into account that certain prognostic parameters are of greater influence than others. Furthermore, many of the available models are partly based on patients collected during the 1980s and even earlier,8,22–24 making it questionable if these are applicable today. Hence there is need for the development of new predictive models, based on more recent study populations and especially with means to take the difference in impact of individual prognostic factors into account.

Objective and aims

The aim of this systematic review was to evaluate currently available prognostic models for the outcome following resection of CRC liver metastases and to assess their differences and applicability in prognosis prediction and decision making.

Materials and methods

Search strategy

A systematic review was performed to assess existing prognostic scoring systems and predictive models considering patients undergoing liver resection for colorectal cancer (CRC) metastases. The Pubmed, Embase and Cochrane Library databases were searched on December 18th 2010, using a syntax including synonyms for ‘colorectal’, ‘liver’, ‘metastasis’, ‘resection’, ‘prognosis’ and ‘prediction’ (Table 1). After removing duplicates, the remaining articles were available for screening.

Article selection

The abstracts of all articles were screened. Animal studies were excluded. For the remaining articles, the inclusion criteria were the mentioning of factors or scoring systems, regarded of prognostic relevance after resection for CRC liver metastases. Those not meeting the inclusion criteria were deemed ‘irrelevant’. Exclusion criteria were studies considering only a certain subgroup, such as reresected patients or major resections alone, reviews, case reports, comments and articles written in a language other than English, German or French.

The selected articles were subjected to a more thorough screening in order to detect articles proposing a new predictive model, risk stratification or classification system. Articles not describing how the prognostic model was developed were excluded. Only articles published in 1990 or later were included. If the same population was used in two different studies, only one was selected. Reference screening of the selected articles was carried out in order to detect any missed articles.

The level of evidence was rated according to the Oxford Centre for Evidence-based Medicine Levels of Evidence.25

Statistical analysis

SPSS software (PASW Statistics 18, SPSS Inc., Chicago, IL, USA) was used to calculate the median and range for study population, follow-up and number of prognostic factors.

Results

Search and selection

The described search yielded 1276 hits in Pubmed, 1586 in Embase and 29 in the Cochrane Library. After removing duplicates a total of 1597 articles remained for screening.

<table>
<thead>
<tr>
<th>Database</th>
<th>Search</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pubmed</td>
<td>(“crc” OR “colorectal”) AND (“liver” OR “hepatic”) AND (“metastasis” OR “metastases” OR “metastatic”) AND (“resection” OR “resections” OR “operation” OR “operations” OR “surgery” OR “surgical”) AND (“prognostic” OR “prognosis” OR “predictive” OR “predict” OR “prediction” OR “risk factor” OR “risk factors” OR “scoring” OR “score” OR “scores” OR “indicator” OR “indicators” OR “model” OR “models”) AND All in [Title/Abstract]</td>
<td>1276</td>
</tr>
<tr>
<td>Embase</td>
<td>The above adjusted for Embase All in :ab,ti</td>
<td>1586</td>
</tr>
<tr>
<td>Cochrane Library</td>
<td>The above adjusted for Cochrane Library All in :ti,ab</td>
<td>29</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2891</td>
</tr>
</tbody>
</table>
The first abstract screening left 302 articles for a more thorough screening, out of which 20 were deemed relevant. However, one study group did not describe how their model, named ‘a preliminary prognostic model’, was developed. Of the remaining 19 articles, 17 were published in 1990 or later, of which two were excluded because of duplicate use of one population. Nagashima et al. created a model in 2004 and then improved this model in 2006 using the same study population. As this was a refinement, only the second study was included. In 2008 Kattan et al. based their model on a population which partly overlapped the one used by Fong et al. in 1999. Since the exact study population was not defined by Kattan et al., only the original model by Fong et al. was included. A final selection of 15 articles remained and reference screening provided no new articles (Fig. 1).

The fifteen proposed models were further evaluated (Table 2). In all studies, Cox proportional hazards regression was used for multivariate analysis to identify the factors that were of prognostic value. Only four studies were prospective. The median population size of the fifteen studies was 305 patients (range 81–1568). Median follow-up time was 32 months (range 16–54 months).

The median number of factors initially analysed per study was 15 (range 4–22) and the median number of factors actually used to base a prognostic model on was 4 (range 2–7). Six studies performed a validation of their own model, either internally (3) or externally, performed on an independent cohort (3). In seven studies a risk categorisation model was developed, in which all factors were weighed equally, whilst the other models assigned different weight to different values, either through a staging system, a formula or different amounts of points assigned to each factor.

All articles were rated, based upon the Oxford Centre for Evidence-based Medicine Levels of Evidence. Because a clinical decision rule was present in all of the studies, the factor that was most important in determining the different levels was the validation of the model on different cohorts, defining most articles as having evidence level 2B and some level 1B.

---

**Figure 1.** Flow chart of systematic selection of articles regarding predictive models for outcome after liver resection for colorectal cancer metastases.

---

*Not meeting the inclusion criteria as mentioned in the text under ‘Article selection’
### Table 2

<table>
<thead>
<tr>
<th>Study characteristics of the evaluated predictive models for outcome after liver resection for colorectal cancer metastases.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study type</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Rees et al.</td>
</tr>
<tr>
<td>Malik et al.</td>
</tr>
<tr>
<td>Minagawa et al.</td>
</tr>
<tr>
<td>Kono et al.</td>
</tr>
<tr>
<td>Nordlinger et al.</td>
</tr>
<tr>
<td>Feng et al.</td>
</tr>
<tr>
<td>Zakaria et al.</td>
</tr>
<tr>
<td>Yamaguchi et al.</td>
</tr>
<tr>
<td>Iwatsuki et al.</td>
</tr>
<tr>
<td>Tan et al.</td>
</tr>
<tr>
<td>Schindl et al.</td>
</tr>
<tr>
<td>Tanaka et al.</td>
</tr>
<tr>
<td>Lise et al.</td>
</tr>
<tr>
<td>Ueno et al.</td>
</tr>
<tr>
<td>Nagashima et al.</td>
</tr>
</tbody>
</table>

Pro = prospective study, Ret = retrospective study, – = unknown, I = internally validated, E = externally validated, N = no validation, OS = overall survival, DSS = disease-specific survival, DFS = disease-free survival, Re = recurrence, R = risk categories with equal weighing of factors, S = unequal weighing of factors through a staging system, O = other way of unequal weighing of factors.

### Prognostic factors

All studies taken together, 25 different factors were identified that independently predicted outcome. However, the number of liver metastases was identified in two more studies (Table 3). The number of liver metastases was identified in both studies with less than five liver metastases. However, the studies applied different cut-off points. The maximum size of liver metastases was identified in five of the studies. Five of them marking a diameter of 5 cm as the cut-off value. Only one study concluded that synchronous metastases presentation was predictive of liver metastases in the group with metachronous presentation. The peripheral factor, which was identified as the most common independent factor, was identified in all studies with a metachronous appearance of CRC and liver metastases. The number of liver metastases was identified as an independent prognostic factor in five studies, five of them marking a diameter of 5 cm as the cut-off value. One study concluded that synchronous metastases presentation was predictive of liver metastases in the group with metachronous presentation. The peripheral factor, which was identified as the most common independent factor, was identified in all studies with a metachronous appearance of CRC and liver metastases. The number of liver metastases was identified as an independent prognostic factor in five studies, five of them marking a diameter of 5 cm as the cut-off value. One study concluded that synchronous metastases presentation was predictive of liver metastases in the group with metachronous presentation.
Table 3  
The occurrence of prognostic factors in the analysis of the selected studies proposing predictive models for outcome after liver resection for colorectal cancer metastases.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of studies (prospective ones)</td>
<td>12 (4)</td>
<td>6086</td>
<td>1B</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Total number of patients</td>
<td>10 (2)</td>
<td>5228</td>
<td>1B</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Highest level of evidence</td>
<td>6 (1)</td>
<td>4340</td>
<td>1B</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Maximum size of metastases</td>
<td>5 (0)</td>
<td>3040</td>
<td>1B</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Interval between CRC operation and detection of liver metastases</td>
<td>5 (3)</td>
<td>2845</td>
<td>1B</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Preoperative CEA-level</td>
<td>5 (1)</td>
<td>2772</td>
<td>1B</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Extrahepatic spread</td>
<td>3 (1)</td>
<td>2311</td>
<td>2B</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Positive resection margin</td>
<td>3 (1)</td>
<td>1439</td>
<td>2B</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Poor differentiation of CRC</td>
<td>3 (1)</td>
<td>1649</td>
<td>1B</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Serosal invasion of CRC</td>
<td>2 (0)</td>
<td>1031</td>
<td>1B</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bilobar spread</td>
<td>2 (0)</td>
<td>454</td>
<td>1B</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

+ = factor identified as predictive, − = factor identified as not predictive, blank field = factor not analysed.
non-anatomical versus anatomical resection,\textsuperscript{36} preoperative glutamic pyruvic transaminase (GPT) level,\textsuperscript{36} serum concentration of alkaline phosphatase and albumine,\textsuperscript{34} tumour doubling time,\textsuperscript{35} marked vascular invasion\textsuperscript{35} and tumour budding\textsuperscript{37} of the primary, perioperative blood transfusion,\textsuperscript{24} postoperative CEA level,\textsuperscript{37} distant lymph node metastases of CRC\textsuperscript{37} and inflammatory response to the tumour (IRT),\textsuperscript{30} which was defined as either a preoperative C-reactive protein (CRP) level greater than 10 mg/L or a neutrophil to lymphocyte ratio greater than 5:1.

\textbf{Discussion}

This review evaluated the currently available prognostic models for predicting outcome following resection of CRC liver metastases. All the included studies deal with prognostic models that are based on risk factors predicting outcome after resection for colorectal cancer liver metastases. Fifteen prognostic systems were identified and there was a tendency towards that the number of liver metastases, CRC spread to lymph nodes, maximum size of metastases, preoperative CEA level, extrahepatic spread and non-radical resection represented independent risk factors.

\textbf{Identified prognostic models}

The studies are not always comparable, since their domain might vary from one including all resected patients\textsuperscript{8,28} to a selected group excluding patients with extrahepatic metastases,\textsuperscript{22,51} a positive margin,\textsuperscript{7,51} or perioperative mortality.\textsuperscript{7,31,37} Furthermore, the primary outcome on which the predictive model was fitted was not overall survival in all cases. Some models did not even have this as an outcome measure, but only regarded disease-specific survival (DSS),\textsuperscript{8} DSS and recurrence,\textsuperscript{24} or disease-free survival (DFS).\textsuperscript{36} Another reason why these studies cannot be directly compared with each other is that they were carried out during different time periods, varying from for example the period 1960—1995\textsuperscript{24} to the period between 1995 and 2005.\textsuperscript{7} During this time new and effective types of management have been introduced and furthermore there has been a change in strategies in the use of chemotherapy in relation to liver surgery, as well as in the indications for surgery.

\textbf{Analysed prognostic factors}

Nonetheless, there is a strong tendency towards the selection of certain risk factors. Nearly all studies regarded the number of liver metastases as a risk factor, with different cut-off points. Apart from these studies, several others have judged this factor to be predictive.\textsuperscript{9,38—42} Studies that supported the contrary are generally a bit older and often based upon smaller study populations.\textsuperscript{17,18,43—46} Spread to lymph nodes by the original CRC is pointed out as a negative prognostic factor by more than half of the reviewed studies and previously other authors came to the same conclusion.\textsuperscript{33,47} However, four reviewed studies found that spread of CRC to lymph nodes was insignificant as a prognostic factor, which is supported by a few older\textsuperscript{45,48,49} and some more recent studies.\textsuperscript{18—20,50} Therefore, this particular predictive factor remains a subject of debate. The same counts for the maximum size of metastases, which was presented as an independent prognostic factor in six of the selected models and many other studies,\textsuperscript{16,17,19,40,44} whilst nine of the selected studies analysed this factor and found it to be insignificant. That conclusion is shared by several other studies.\textsuperscript{18,20,49—51}

The interval between the CRC operation and detection of liver metastases was analysed as a factor in only eight studies, out of which five deemed it predictive. There are some studies both supporting\textsuperscript{16,38,48} and contradicting\textsuperscript{18,20,52} this. The difference between synchronously and metachronously presenting metastases was investigated by ten groups, nine judging it to lack prognostic value, which has been confirmed by many others in the past.\textsuperscript{16,18—20,48,51,53,54}

A factor that often has been deemed a negative prognostic factor is the preoperative CEA level.\textsuperscript{18,20,40,50,55,56} Of the analysed prognostic models, five also found this to be predictive, whilst seven studies did not find this significant. All in all, the predictive value of preoperative CEA level remains arbitrary, partly because different cut-off values were used in different studies, although it seems to potentially possess predictive relevance when applying a high cut-off value.

Since patients with extrahepatic spread and a positive resection margin were excluded from nearly all analysed studies, the prognostic value of these factors was only investigated in seven and five studies, respectively. Most authors valued them as negative predictive factors and it was often regarded as a contraindication in the first place. Both previous and recent literature confirm a positive resection margin as a negative prognostic factor.\textsuperscript{16,18,19,42,45,51,52,57} Extrahepatic spread has also been regarded as such,\textsuperscript{17,55} but some have proven the opposite\textsuperscript{16,15} and it is especially important to realise that e.g. lung metastases may not represent a contraindication for liver resection and long-term survival can still be achieved in case these other metastases also can be removed completely.\textsuperscript{58,59} Ten of the studies found bilobar spread not to represent a prognostic factor, which is also supported by many other studies outside of this review.\textsuperscript{16,17,20,39,45,48}

Some factors that are currently known to be of prognostic value were not analysed by any of the mentioned models. An example is neoadjuvant chemotherapy before liver resection. The administration of preoperative chemotherapy has been shown to improve the postoperative disease-free survival.\textsuperscript{60} It has also been shown that progress of disease during administration of neoadjuvant chemotherapy before liver resection is a negative prognostic factor both for overall for survival and disease-free survival.\textsuperscript{51,62} This seems even more evident when
specifically analysing the histopathological response to chemotherapy, some of which has been identified as a prognostic factor affecting survival. Since both of these findings result from more recent studies, it can be explained that none of the studied prognostic models made use of them, as they are based on older study groups. However, preoperative CEA-levels and tumour size have been identified to be predictive for the pathologic response, suggesting a relation between these older and more recently identified prognostic factors.

**Validation**

In order to test the applicability of a model, one could ideally test the model on an independent data set from another institution (external validation) or on a second data set from the same centre (temporal validation). However, since these possibilities are often not at hand, internal validation is frequently used as an alternative, for example by means of bootstrapping. Just six of the selected study groups performed validation. Rees et al., Konopke et al., and Nordlinger et al. used the bootstrapping method, whilst external validation was performed by Minagawa et al., Schindl et al., and Nagashima et al. Nonetheless, some models have been validated externally by other study groups. Especially the model created by Fong et al. has been validated by other study groups.

**Applicability**

Another point of discussion is the variation in study methods and types of models that were generated by the selected study groups, and thus their applicability. In terms of the identification of prognostic factors, most studies had the disadvantage of being retrospective. Also, the study populations varied a lot in size, from 81 patients to 1568 patients. The liability of the largest study, performed by Nordlinger et al., may be questioned. It was carried out by sending questionnaires to many institutions. A selection bias is present in the response rate of the institutions and it is doubtful whether the information regarding diagnostics, treatment and follow-up were actually comparable between all 85 hospitals.

In many of the predictive models, the identified risk factors were all weighed equally, for example by awarding one point to the presence of each risk factor. However, Rees et al. gave the risk factors an individual number of points according to their influence on prognosis and Minagawa et al. selected one of their four factors as a predictor of most severe prognosis. Yamaguchi et al., Ueno et al., and Tanaka et al. all had a slightly more complicated staging system, which also considered factors with an intermediate influence level. Iwatsuki et al., Schindl et al., and Nagashima et al. used the hazard ratio for the different risk factors.

**Future aspects**

To do justice to possible influence of separate factors on each other and the value of their combination, a predictive model using artificial neural networks (ANN) could be effective. An ANN is a mathematical and computational method, which resembles the structure and function of biological neural networks, such as in the human brain. ANNs consist of processing units (nodes), which are highly interconnected via a set of weights, as in a network of neurons, and can process data in a manner analogous to the brain. By training the neural network, both linear and non-linear relationships between input and output variables can be discovered.

ANN has already been applied successfully in the clinical setting, for example in diagnosing colorectal cancer, in the prediction of outcome after acute pancreatitis and mortality from liver cirrhosis, for differentiation between liver lesions on imaging and for mortality risk scoring in cardiac surgery. So far, no one has applied this system to predict the prognosis after liver resection for colorectal cancer metastases. With regard to the previously mentioned advantages, this could be an attractive possibility for the future, especially since we most probably will be able to analyse more complex prognostic factors, including the pathological and clinical response to more advanced oncological and surgical treatment strategies.

**Conclusion**

Currently there are several different prognostic models to predict survival or recurrence after liver resection for colorectal cancer metastases. However, the identified factors vary considerably between models and only few systems take into account that different factors have different levels of prognostic significance. Artificial neural networks have not yet been applied to predict outcome for this patient group and would be a potentially interesting step forward in the future creation of a prognostic model.

**Conflict of interest statement**

There are no conflicts of interest to be declared. No financial support was requested or received in the production of this article.

**References**


