ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

Σχολή Επιστημών Υγείας

Τμήμα Ιατρικής

PROGNOSTIC SYSTEMS AND THE NATURAL HISTORY OF CIRRHOSIS

Ιωάννα Α. Μοσχανδρέα

Διδακτορική Διατριβή

Ηράκλειο, 2002

Επιβλέπων Καθηγητής

Καθηγητής ……..……...…...Η. Κουρούμαλης

Συμβουλευτική Επιτροπή

Καθηγητής ……..……...…...Η. Κουρούμαλης

Καθηγητής ……..……...……..Γ. Ηλιόπουλος

Επίκ. Καθηγητής ……..……...…...Ι. Μουζάς

Εξεταστική επιτροπή

Καθηγητής ……..……...……....Η. Κουρούμαλης

Καθηγητής ……..……...…..…….Γ. Ηλιόπουλος

Καθηγητής ……..……...……....Α. Καφάτος

Επίκ. Καθηγητής ……..……...…...Μ. Αλεξανδράκης

Επίκ. Καθηγητής ……..……...……....Ι. Διαμαντής

Επίκ. Καθηγητής ……..……...…...Γ. Κολιός

Επίκ. Καθηγητής ……..……...…...Ι. Μουζάς

Αφιερώνεται στους γονείς μου Μαρία και Ανδρέα

Acknowledgements

It is the unfailing encouragement and support of Professor Elias Kouroumalis, Director of the University Hospital Dept. of Gastroenterology, that led to the undertaking and successful completion of this PhD thesis. A simple "thank you" would not be enough to express my gratitude. Many thanks go to the clinicians at the Gastroenterology Clinic, for our successful collaboration in relevant publications and also for data collection. Ass. Professor Michael Alexandrakis of the Heamatology Dept. I thank for his help and cooperation as regards the ascitic patient data. My special thanks are extended to Dr Panayiotis Skordilis, not only for his enthusiastic collaboration in publications but also for going out of his way to help me on multiple occasions when the going got tough!

I am indebted to Professor Anthony Kafatos for his support in this and my other academic endeavours. The Preventive Medicine & Nutrition research team has provided a friendly and conducive working atmosphere over the last five years. Thanks to all. Others at the Medical School have also supported my efforts – thanks in particular to Mr Takis Efthymiou.

I would like to thank my partner, Yiannis Tselos, for his patience, and more importantly, for his understanding. My family has always been my guiding force. My parents have been a constant source of strength for me throughout my academic progression, offering both moral and financial support. They, and my godfather Nikos Vardis, have always been there to turn to for advice and encouragement. Finally, the impending arrival of my son Alexander Polivios was a significant contributory factor for the prompt completion of this thesis!

Περίληψη της Διατριβής

ΠΡΟΓΝΩΣΤΙΚΑ ΣΥΣΤΗΜΑΤΑ ΚΑΙ ΦΥΣΙΚΗ ΠΟΡΕΙΑ ΤΗΣ ΚΙΡΡΩΣΕΩΣ.

Εισαγωγή

Η κίρρωση του ήπατος οφείλεται σε διάφορες αιτίες. Οι αιτίες περιλαμβάνουν τον αλκοολισμό, τις χρόνιες ιογενείς ηπατίτιδες Β, C και D, την αυτοάνοσο ηπατίτιδα, τις κληρονομικές νόσους, την μη-αλκοολική στεατοηπατίτιδα, τα χολοστατικά νοσήματα (π.χ. πρωτοπαθή χολική κίρρωση), την χρήση ναρκωτικών (μέσω ιώσεων), την επαφή με τοξίνες, τις λοιμώξεις και άλλες. Η πιο συχνή αιτία της κιρρώσεως παγκοσμίως θεωρείται η ηπατίτιδα Β, ενώ στις Ηνωμένες Πολιτείες οι συχνότερες αιτίες είναι ο χρόνιος αλκοολισμός και η ηπατίτιδα C. Σχετικά στοιχεία για την Ελλάδα δεν υπάρχουν. Οι πιθανές επιπλοκές της κιρρώσεως είναι αρκετές, από τις οποίες οι πιο συχνές είναι ο ασκίτης, η ηπατική εγκεφαλοπάθεια και η κιρσορραγία. Το ηπατοκυτταρικό καρκίνωμα (ΗΚΚ) μπορεί να δημιουργηθεί ως επιπλοκή της κιρρώσεως. Είναι γνωστό ότι η χρόνια προσβολή από τον ιό της ηπατίτιδας Β αποτελεί σοβαρό παράγοντα κινδύνου για την εμφάνιση ηπατοκυτταρικού καρκινώματος ενώ τα τελευταία έτη υπάρχουν ενδείξεις ότι και ο ιός της ηπατίτιδας C είναι εξίσου σημαντικός παράγοντας κινδύνου. Στην Κρήτη, ο επιπολασμός της ηπατίτιδας Β (HBV) είναι πολύ χαμηλότερος απ' ότι στην ηπειρωτική Ελλάδα ενώ ο επιπολασμός της ηπατίτιδας C (HCV) είναι το κύριο πρόβλημα στον άνω τον 40 ετών πληθυσμό της Κρήτης.

Ο στόχος της παρούσας εργασίας είναι η μελέτη των κιρρωτικών νόσων στην Κρήτη. Στη μελέτη της κιρρώσεως, όπως και άλλων ασθενειών, μεγάλο ενδιαφέρον παρουσιάζουν τα προγνωστικά μοντέλα του χρόνου επιβίωσης και η ταξινόμησης ασθενών σε ομάδες με διαφορετικές προγνώσεις οι οποίες επηρεάζουν τη θεραπευτική αγωγή. Ο κύριος σκοπός της παρούσας μελέτης είναι συνεπώς η δημιουργία και επικύρωση προγνωστικών μοντέλων, όπως μοντέλων χρόνου επιβίωσης, μοντέλων διάκρισης ομάδων ασθενών, μοντέλων επίπτωσης του ηπατοκυτταρικού καρκινώματος κι άλλων. Οι σχετικές τεχνικές εφαρμόσθηκαν στις ακόλουθες ομάδες ασθενών:

α) ασθενείς οι οποίοι διεγνώσθησαν στη κλινική με κίρρωση: η φυσική πορεία της κιρρώσεως στη Κρήτη ανά αιτιολογία

β) κιρρωτικοί ασθενείς με HCV υπό θεραπεία με *Plaquenil*.

γ) ασθενείς με πρωτοπαθή χολική κίρρωση (ΠΧΚ) υπό θεραπεία με UDCA: συγκρίσεις με το προγνωστικό μοντέλο Mayo.

δ) ασθενείς στους οποίους εμφανίστηκε ΗΚΚ: η φυσική πορεία του ΗΚΚ

ε) ασθενείς με ΗΚΚ υπό θεραπευτική αγωγή με **οκτρεοτίδη** (octreotide): μία τυχαιοποιημένη, ελεγχόμενη μελέτη ασθενών υπό θεραπεία με οκτρεοτίδη έναντι ασθενών άνευ θεραπείας.

στ) ασθενείς με ΗΚΚ υπό θεραπεία με **σωματοστατίνη μακράς δράσης** (long-acting somatostatin analogue): μία μελέτη ασθενών με ΗΚΚ υπό θεραπευτική αγωγή σε σύγκριση με ομάδα ιστορικού ελέγχου.

η) ασθενείς που παρουσιάζουν ασκίτη: ι) διάκριση μεταξύ κακοήθους και μη-κακοήθους κιρρωτικού ασκίτη ιι) διάκριση μεταξύ τριών μορφών περιτονικής διάχυσης (peritoneal effusion) σε ασθενείς με ασκίτη.

Επί πλέον, παρουσιάζονται τα αποτελέσματα τριών επιδημιολογικών επισκοπήσεων σχετικά με τον επιπολασμό των ιολογικών δεικτών HBV και HCV στην Κρήτη.

Τα στοιχεία που έχουν χρησιμοποιηθεί περιλαμβάνουν όλους τους ασθενείς με κίρρωση πού έχουν νοσηλευτεί στο ΠεΠαΓΝΗ από την έναρξη λειτουργίας της Γαστρεντερολογικής Κλινικής μέχρι τον Οκτώβρη του 2000. Συνολικά, 470 ασθενείς κατεγράφησαν σε βάση δεδομένων, έχοντας εισέλθει στην κλινική με κίρρωση ή κάποια επιπλοκή της κιρρώσεως σε αυτό το διάστημα. Οι 139 (29%) παρουσιάστηκαν στην κλινική με ρήξη της αντισταθμίσεως. Επίσης, από τον Σεπτέμβριο του 1989 μέχρι τον Μάρτιο του 2000, διεγνώσθησαν 114 άτομα με ΠΧΚ, τα οποία έλαβαν θεραπεία με UDCA.

Μέθοδοι

Α) Για την κάθε ομάδα ασθενών στην οποία ο στόχος ήταν η εκτίμηση των χρόνων επιβίωσης, εφαρμόστηκε αρχικά η μη-παραμετρική μέθοδος Kaplan-Meier για σχεδιασμό καμπυλών επιβίωσης για κάθε παράγοντα, με μηδενική υπόθεση ότι ο κίνδυνος θανάτου σε κάθε χρονική στιγμή είναι ο ίδιος για όλα τα επίπεδα του εκάστοτε παράγοντα. Ακολούθησε πολυπαραγοντική ανάλυση τύπου Cox, χρησιμοποιώντας διαδικασίες επιλογών βήμα-προς-βήμα (stepwise selection procedures) για την εύρεση των σημαντικών μεταβλητών, δηλαδή των μεταβλητών που επηρεάζουν τον χρόνο επιβίωσης. Τα μοντέλα Cox έχουν τη μορφή

$h_i(t) = h_0(t)e^{\sum_{j=1}^{p}\beta_j x_{ji}}$ όπου το x_{ji} είναι η τιμή της ανεξάρτητης μεταβλητής X_j, j=1,..,p για τον ασθενή i, (i=1,…,n), το $h_i(t)$ είναι η συνάρτηση επικινδυνότητας (hazard function) του ασθενή i, και το $h_0(t)$ είναι η βασική συνάρτηση επικινδυνότητας. Η συνάρτηση επικινδυνότητας δείχνει τον κίνδυνο θανάτου κάθε στιγμής υπό την προϋπόθεση ότι δεν έχει συμβεί ο θάνατος μέχρι την συγκεκριμένη στιγμή. Όπως φαίνεται από την παραπάνω εξίσωση, στο μοντέλο Cox (όπου δεν γίνεται καμία υπόθεση για την μορφή που λαμβάνει η βασική συνάρτηση επικινδυνότητας) ο λόγος του $h_i(t)$ για δύο άτομα παραμένει σταθερός στην διάρκεια του χρόνου. Οι μέθοδοι Kaplan-Meier και Cox έχουν το πλεονέκτημα (εν σχέσει με τις συνήθεις μεθόδους όπως η λογιστική παλινδρόμηση) ότι λαμβάνουν υπ' όψιν όχι μόνο το τελικό αποτέλεσμα, αλλά και τον χρόνο παρακολούθησης, συμπεριλαμβάνοντας τα άτομα τα οποία δεν παρακολουθήθηκαν μέχρι το τελικός, και θεωρείται ότι έχουν υποστεί «λογοκρισία» (censoring).

Στα δεδομένα των κιρρωτικών, μερικές από τις ανεξάρτητες μεταβλητές μπορούν να θεωρηθούν ως χρονο-εξαρτημένες, όπως η μεταβλητή που περιγράφει το αν υπήρξε ρήξη και αυτή που περιγράφει την ύπαρξη ή μη του ΗΚΚ. Οι χρονο-εξαρτημένες μεταβλητές περιελήφθησαν στα σχετικά μοντέλα. Για τη διαπίστωση τυχόν έλλειψης καλής προσαρμογής του μοντέλου επιβίωσης και για τον εντοπισμό ακραίων τιμών (outliers) υπολογίσθηκαν οι ακόλουθοι τρεις τύποι υπολοίπων (residuals): υπόλοιπα Cox-Snell, υπόλοιπα martingale και μερικά υπόλοιπα. Οι τεχνικές bootstrap και jackknife χρησιμοποιήθηκαν στη παρούσα μελέτη για να ερευνηθεί η σταθερότητα των παλινδρομικών μοντέλων Cox σε όρους επιλογής των μεταβλητών που περιλαμβάνονται στα μοντέλα επιβίωσης και στα μοντέλα του χρόνου μέχρι τη ρήξη της αντισταθμίσεως. Η εμπειρική βαθμολόγηση Brier και οι εκτιμήσεις της εξηγηθείσας υπολοιπόμενης μεταβλητότητας (explained residual variation) υπολογίσθηκαν για τα μοντέλα πρόβλεψης του χρόνου μέχρι τη ρήξη. Συνεπώς αξιολογήθηκε η ακρίβεια των συγκεκριμένων μοντέλων Cox σε σχέση με τα αντίστοιχα απλά μοντέλα Kaplan-Meier.

Οι αδροί ρυθμοί επίπτωσης του ΗΚΚ στους Κρήτες κιρρωτικούς εκτιμήθηκαν, εκφραζόμενοι ως αριθμός ανθρωπο-ετών παρακολούθησης. Έγιναν επίσης εκτιμήσεις

των αθροιστικών ρυθμών επίπτωσης του ΗΚΚ χρησιμοποιώντας το μοντέλο παλινδρομικής ανάλυσης αναλογικών κινδύνων του Cox.

Β) Για την ομάδα ασθενών με ΠΧΚ στους οποίους χορηγήθηκε η θεραπευτική αγωγή UDCA, έγινε σύγκριση του χρόνου επιβίωσης με τον προβλεπόμενο χρόνο επιβίωσης βάσει του προγνωστικού μοντέλου «Mayo». Το μοντέλο Mayo είναι βασισμένο σε ασθενείς με ΠΧΚ που δεν ακολούθησαν θεραπευτική αγωγή και επομένως περιγράφει την φυσική πορεία της ΠΧΚ. Χρησιμοποιώντας το «Mayo risk score» (ο δείκτης κινδύνου με βάση το μοντέλο Mayo, ο οποίος βασίζεται σε συνδυασμό 5 μεταβλητών) εκτιμήθηκε καμπύλη επιβίωσης για τα πρώτα 7 έτη μετά την διάγνωση για τον κάθε ασθενή. Έγιναν γραφικές συγκρίσεις της εκτιμώμενης καμπύλης επιβίωσης Kaplan-Meier με τη προβλεπόμενη καμπύλη που προέκυψε από την εφαρμογή του μοντέλου Mayo. Για την κατασκευή της δεύτερης καμπύλης χρησιμοποιήθηκε η απευθείας-προσαρμοζόμενη μέθοδος (direct-adjusted method). Η δοκιμασία μοναδικού-δείγματος log rank (one-sample log-rank test) εφαρμόστηκε για να διαπιστωθούν τυχόν διαφορές μεταξύ της προβλεπόμενης επιβίωσης ασθενών άνευ θεραπείας (του μοντέλου Mayo) και της επιβίωσης των ασθενών στους οποίους χορηγήθηκε η θεραπευτική αγωγή UDCA.

Γ) Δύο μέθοδοι πολυμεταβλητής ανάλυσης εφαρμόστηκαν για τη διάκριση μεταξύ ομάδων ασκιτικών. Ο σκοπός ήταν να ελεγχθεί αν η κατάταξή τους ανάλογα με την κλινική τους διάγνωση μπορεί να προβλεφθεί από τις βιοχημικές τους μετρήσεις. Υπετέθη ότι η διάγνωση έγινε ανεξάρτητα από τις βιοχημικές μετρήσεις που περιελήφθησαν στα μοντέλα. Η πρώτη μέθοδος ανάλυσης λέγεται αναδρομικός διαμερισμός (recursive partitioning). Σε αυτή τη τεχνική χρησιμοποιείται ένας αλγόριθμος δυαδικού διαμερισμού (binary partitioning algorithm) που χωρίζει τις συμμεταβλητές στο σημείο το οποίο έχει κριθεί, βάσει του αλγορίθμου, ως το πιο σημαντικό. Με αυτό τον τρόπο δημιουργείται ένα σύνολο δυαδικών μεταβλητών. Χρησιμοποιήθηκε ένα μέτρο ετερογενότητας σε κάθε <κόμβο>, δηλαδή σε κάθε σημείο που γίνεται ο χωρισμός των μεταβλητών. Η πρώτη διαχώριση (split) παρέχει την καλύτερη πρόβλεψη των ομάδων. Η δεύτερη μέθοδος που εφαρμόστηκε ήταν η διακρίνουσα ανάλυση (discriminant analysis). Έγινε απεικόνιση του διαχωρισμού των ομάδων με βάση τις δυο πρώτες διακρίνουσες συναρτήσεις. Για να ερευνηθεί η

δυνατότητα εφαρμογής του τελικού μοντέλου σε νέες περιπτώσεις, χρησιμοποιήθηκε η μέθοδος της διασταυρωτικής επικύρωσης (cross-validation).

Αποτελέσματα

Φυσική πορεία της κιρρώσεως

Από τους 312 ασθενείς που διεγνώσθησαν με αντιρροπούμενη κίρρωση, οι 169 (54%) ήταν άνδρες ενώ από τους 138 ασθενείς που παρουσιάσθηκαν με μη-αντιρροπούμενη κίρρωση οι 107 (78%) ήταν άνδρες. Οι 154 από τους ασθενείς με αντιρροπούμενη κίρρωση (49%) έπαθαν ρήξη της αντισταθμίσεως στην διάρκεια της παρακολούθησής τους. Η εκτιμώμενη διάμεσος του χρόνου μέχρι την ρήξη της αντισταθμίσεως ήταν 58 μήνες (95% Δ.Ε. 51 με 65 μήνες) αλλά βρέθηκε ότι διαφέρει σημαντικά ανάλογα με την αιτία της κιρρώσεως (log rank test, $p<0.0001$)˙ ήταν 81 μήνες στους 145 ασθενείς (47%) των οποίων η κίρρωση προήλθε από τον ιό της ηπατίτιδος C (95% Δ.Ε. 45 με 117 μήνες) ενώ ήταν μόνο 35 μήνες (95% Δ.Ε. 19 με 51 μήνες) στους 56 (18%) ασθενείς με αιτιολογία τον αλκοολισμό και 36 μήνες (95% Δ.Ε. 20 με 52 μήνες) στους 45 ασθενείς (15%) των οποίων η κίρρωση προήλθε από τον ιό της ηπατίτιδος Β. Οι 17 ασθενείς (6%) των οποίων το αίτιο ήταν συνδυασμός του ιού Β ή C με τον αλκοολισμό είχαν συνολικό διάμεσο χρόνο ρήξεως της αντισταθμίσεως μόνο 31 μήνες (95% Δ.Ε. 16 με 46 μήνες). Το ποσοστό των κιρρωτικών ασθενών οι οποίοι παρέμειναν ελεύθεροι-ρήξης (decompensation-free) τρία έτη μετά την διάγνωση εκτιμήθηκε να είναι 65% (95% Δ.Ε. 60% με 71%) και μετά από επτά έτη 34% (95% Δ.Ε. 26% με 42%).

Εφαρμόζοντας το μοντέλο Cox για το χρόνο μέχρι τη ρήξη της αντισταθμίσεως, βρέθηκε ότι οι σημαντικοί προγνωστικοί παράγοντες ήταν η ηλίκια κατά τη διάγνωση (σε έτη, Σ.Κ. 1,02) και η αιτιολογία της κιρρώσεως (Σ.Κ. 0,58 για τους ασθενείς με HCV σε σχέση με τους ασθενείς με κρυψιγενή αιτιολογία κιρρώσεως ενώ ο αντίστοιχος Σ.Κ. για αυτούς με αιτιολογία τον αλκοολισμό ήταν 1,72). Η ανάλυση bootstrap χρησιμοποιήθηκε για την επιβεβαίωση της σταθερότητας του μοντέλου. Επίσης δημιουργήθηκε ένας δείκτης πρόγνωσης (ΔΠ) της ρήξης της αντισταθμίσεως όπου

ΔΠ=0,016 * (ηλικία –62,29) + 0,54*Δκτ{αλκοολισμός} + 0,40*Δκτ(HBV}-
0,54*Δκτ{HCV} + 0,45 *Δκτ{αλκοολισμός + ηπατίτιδα}

με Δκτ{Χ}=1 όταν Χ είναι η αιτιολογία, αλλιώς είναι 0. Οι πιθανότητες του να παραμένει ένας ασθενείς ελεύθερος-ρήξης στα 3-, 5-και 7- έτη μετά τη διάγνωση, δεδομένης της τιμής του ΔΠ του συγκεκριμένου ασθενούς, υπολογίσθηκαν και απεικονίσθηκαν γραφικώς.

Η εκτιμώμενη διάμεσος του γενικού χρόνου από την ρήξη της αντισταθμίσεως μέχρι τον θάνατο ήταν 59 μήνες (95% Δ.Ε. 43 με 76 μήνες) . Εφαρμόζοντας μοντέλο τύπου Cox με αρχικές προγνωστικές μεταβλητές το φύλο, την ηλικία, το αίτιο της κιρρώσεως και την μορφή της ρήξης, με την διαδικασία των βήμα-προς-βήμα επιλογών βρέθηκε ότι μόνο η μορφή της ρήξης και η ηλικία επηρέαζαν σημαντικά τον χρόνο επιβίωσης˙ ο Σ.Κ. θανάτου ήταν 5,7 φορές υψηλότερος στα άτομα των οποίων η ρήξη ήταν εγκεφαλοπάθεια, εν σχέσει με τα άτομα που παρουσιάστηκαν με κιρρσοραγία.

Οι 70 από τους 312 αντιρροπούμενους κιρρωτικούς απεβίωσαν από ηπατική ανεπάρκεια κατά τη διάρκεια της μελέτης. Η διάμεσος του χρόνου επιβίωσης ήταν 126 μήνες (με 95% Δ.Ε. 103 έως 149 μήνες). Όπως παρατηρήθηκε και στο μοντέλο του χρόνου μέχρι τη ρήξη, οι αντιρροπούμενοι ασθενείς με HCV φάνηκε να έχουν κατά μέσο όρο μεγαλύτερο χρόνο επιβίωσης από ότι οι ασθενείς των άλλων αιτιολογικών ομάδων (log rank p=0,0024). Ο Σ.Κ. για τους κιρρωτικούς με αιτιολογία την HCV ήταν 0,34 σε σχέση με ασθενείς με κρυψιγενείς αιτιολογίες (95% Δ.Ε. 0,17 με 0,69). Οι σημαντικοί προγνωστικοί παράγοντες του μοντέλου Cox ήταν το φύλο, η ηλίκια και η αιτιολογία της κιρρώσεως. Όταν συμπεριελήφθησαν στο μοντέλο οι χρόνο-εξαρτημένες δυαδικές μεταβλητές που αναφέρονται στην ύπαρξη της ρήξης και την ύπαρξη του ΗΚΚ, βρέθηκαν να είναι οι μόνες σημαντικές μεταβλητές της πρόβλεψης του χρόνου επιβίωσης

Από τα 410 άτομα που διεγνώσθησαν με κίρρωση (αντιρροπούμενη ή μη-), τα 39 (9,5%) προσεβλήθησαν από ΗΚΚ μετά τη διάγνωση της κιρρώσεως (υπήρχαν και 6 επιπλέον άτομα που προσεβλήθησαν από ΗΚΚ μέσα σε ένα μήνα από την διάγνωση). Ο αδρός ρυθμός επίπτωσης του ΗΚΚ στους κιρρωτικούς ασθενείς που διεγνώσθησαν στην κλινική εκτιμήθηκε ότι είναι 2,3 ανά 100 ανθρωπο-έτη και ο γενικός διάμεσος χρόνος μέχρι την εμφάνιση του ΗΚΚ 10 έτη και ένας μήνας (95% Δ.Ε. από 9 έτη και 8 μήνες μέχρι 10 έτη και 7 μήνες). Οι γενικοί αθροιστικοί ρυθμοί

επίπτωσης του ΗΚΚ εκτιμήθηκαν να είναι 8% στα 3 έτη και 15% στα 5 έτη μετά τη διάγνωση. Για τους ασθενείς με αιτιολογία την HCV, οι εκτιμώμενοι αθροιστικοί ρυθμοί επίπτωσης του ΗΚΚ ήταν 7% στα 3 έτη και 9% στα 5 έτη μετά τη διάγνωση ενώ τα αντίστοιχα ποσοστά για ασθενείς με αιτιολογία την HBV ήταν 20% και 27% αντιστοίχως.

Συγκρίσεις του χρόνου επιβίωσης των ασθενών με ΠΧΚ με το διεθνές προγνωστικό μοντέλο Mayo

Το 89% των ασθενών με ΠΧΚ ήταν γυναίκες. Τα δεδομένα είχαν υποστεί «βαριά λογοκρισία» (heavily censored) διότι μόνο 17 άτομα απεβίωσαν κατά το χρόνο παρακολούθησης οπότε οι εκτιμήσεις του χρόνου επιβίωσης δεν είναι τόσο σίγουρες. Η διάμεσος του χρόνου επιβίωσης εκτιμήθηκε ότι είναι 117 μήνες (με 95% Δ.Ε. από 107 έως 127 μήνες). Οι πιθανότητες επιβίωσης άνευ αγωγής υπολογισμένες με το μοντέλο Mayo (simulated controls) ήταν σημαντικά χαμηλότερες από αυτές που εκτιμήθηκαν βάσει της μεθόδου Kaplan-Meier (χ^2=12.81 με 1β.ε., p<0.001, προβλεπόμενος αριθμός θανάτων = 39). Η διαφορά μεταξύ του παρατηρηθέντος και του αναμενόμενου χρόνου επιβίωσης φαίνεται να αυξάνεται καθώς ο χρόνος από τη διάγνωση αυξάνεται. π.χ. 1 έτος μετά τη διάγνωση τα παρτηρηθέντα και αναμενόμενα ποσοστά επιβίωσης συμπίπτουν (στο 95%) ενώ στα 7 έτη το παρατηρηθέν ποσοστό επιβίωσης είναι 82% και το αναμενόμενο είναι 62%. Ένα πιθανό συμπέρασμα είναι ότι η θεραπευτική αγωγή παρατείνει την ζωή των ασθενών, και πιο ειδικά των ασθενών οι οποίοι δεν έχουν τη βαριά μορφή της νόσου.

Πρόγνωση του ηπατοκυτταρικού καρκινώματος.

Μια ομάδα ασθενών στην οποία έγινε εκτίμηση της φυσικής πορείας της νόσου ήταν τα 73 άτομα που διεγνώσθησαν με ΗΚΚ στο ΠεΠαΓΝΗ μεταξύ 1992 και 1996. Οι ασθενείς ήταν ως επί το πλείστον άνδρες (84%) και κιρρωτικοί (85%). Από την ανάλυση προέκυψαν αποδείξεις ότι υπάρχουν σημαντικές διαφορές στο χρόνο επιβίωσης των ασθενών με ΗΚΚ, οι οποίες εξαρτώνται από το μέγεθος του όγκου, το στάδιο κατά Okuda, την παρουσία HBeAg (θετικό/ αρνητικό), την συγκέντρωση λευκωματίνης και το anti-HBc (θετικό / αρνητικό).

Το 1996 έγινε σύγκριση στον χρόνο επιβίωσης 28 ασθενών με ΗΚΚ που έλαβαν μια καινούργια θεραπεία (octreotide) με ασθενείς που δεν έλαβαν θεραπεία (30 άτομα). Η

κλινική δοκιμή που είχε διάρκεια 4 ετών ήταν τυχαιοποιημένη και ελεγχόμενη. Οι ασθενείς είχαν παρόμοιους προγνωστικούς παράγοντες (ηλικία, φύλο, Okuda, ύπαρξη κιρρώσεως, ιολογική εικόνα, βιοχημικές μετρήσεις). Βρέθηκε ότι η διάμεσος του χρόνου επιβίωσης ήταν 13 μήνες (95% Δ.Ε. 9-17 μήνες) για τους ασθενείς που είχαν θεραπευτική αγωγή και 4 μήνες (95% Δ.Ε. 2-6 μήνες) για τους ασθενείς που δεν είχαν θεραπευτική αγωγή. Χρησιμοποιώντας ένα μοντέλο τύπου Cox βρέθηκε ότι, ανεξάρτητα από την επίδραση της θεραπείας στον χρόνο επιβίωσης, υπήρχε και θετική επίδραση της συγκέντρωσης λευκωματίνης και αρνητική επίδραση από την παρουσία κιρρώσεως. Οι κλινικές δοκιμές αυτής της μορφής δεν συνεχίστηκαν διότι θεωρήθηκε πλέον ότι δεν ήταν ηθικά επιτρεπτές εφόσον η θεραπεία φαίνεται να παίζει τόσο σημαντικό ρόλο στην παράταση της ζωής, και μάλιστα με πολύ λίγες αρνητικές παρενέργειες. Κατά συνέπεια, η μεταγενέστερη εκτίμηση θεραπείας πραγματοποιήθηκε χρησιμοποιώντας ιστορική ομάδα ελέγχου (historical controls). Τα αποτελέσματα, όσον αφορά διαφορές στον χρόνο επιβίωσης μεταξύ των δυο ομάδων, ήταν παρόμοια με αυτά της προηγούμενης μελέτης. Η εκτιμηθείσα συνάρτηση επικινδυνότητας για τους ασθενείς που έλαβαν θεραπεία ήταν σε επίπεδο 0,37 εν σχέσει με την ομάδα ελέγχου (95%Δ.Ε. 0,18 έως 0,73), έχοντας λάβει υπ' όψιν πιθανές διαφορές σε άλλους προγνωστικούς παράγοντες. Επειδή θεωρήθηκε ότι μπορεί να υπήρξε μεροληψία λόγω του ότι η ομάδα ελέγχου ήταν ιστορική, έγινε επανάληψη της ανάλυσης αφού πρώτα αφαιρέθηκαν από την ιστορική ομάδα οι 5 ασθενείς με χρόνο επιβίωσης λιγότερο των 5 μηνών. Οι ασθενείς που έλαβαν την θεραπεία παρουσίασαν και πάλι σημαντικά αυξημένο χρόνο επιβίωσης, όπως προέκυψε τόσο από το μονομεταβλητό μοντέλο όσο και από το πολυμεταβλητό μοντέλο τύπου Cox.

Μαθηματικά μοντέλα για την διαχώριση των ασθενών ανάλογα με την φύση του ασκίτη

Με την μέθοδο του αναδρομικού διαμερισμού και χρησιμοποιώντας το λόγο της μέτρησης στο ασκιτικό υγρό με αυτόν στον ορό για διάφορες πρωτεΐνες (πρωτεΐνες οξείας φάσεως, ανασοσφαιρίνες and κυτταροκίνες) και συμπληρώνοντάς τον με άλλες βιοχημικές μετρήσεις όπως της ολικής πρωτεΐνης, της λευκωματίνης και της LDH, πραγματοποιήθηκε η κατάταξη των ασθενών σε εκείνους με κακοήθη ασκίτη και εκείνους με ασκίτη προερχόμενο από κίρρωση, με επιτυχία 100%. Στο μοντέλο χρησιμοποιήθηκαν μόνο οι μετρήσεις της λευκωματίνης και της IL-1α στο ασκιτικό

υγρό. Οι ‘κανόνες κατάταξης’ που δημιουργήθηκαν προβλέπουν ότι μια μέτρηση του λόγου της λευκωματίνης κάτω από 0,39 υποδηλώνει την ύπαρξη κίρρωσης και όχι νεοπλάσματος ενώ μια μέτρηση του λόγου της λευκωματίνης πάνω από 0,39 σε συνδυασμό με λόγο της IL-1α κάτω από 2,17 υποδηλώνει την ύπαρξη νεοπλάσματος. Αν η τελευταία μέτρηση είναι άνω των 2,17 ο κίνδυνος μειώνεται στο περίπου 40% ότι υπάρχει νεόπλασμα. Επειδή όμως οι ασθενείς στην τελευταία κατηγορία (δηλαδή με λόγο της IL-1α πάνω από 2,17) ήταν μόνο πέντε, δεν μπορούμε να θεωρήσουμε ως βέβαιη την αξιοπιστία του συγκεκριμένου διαχωρισμού.

Με την μέθοδο της διακρίνουσας ανάλυσης για την κατάταξη σε μια από τις τρεις ομάδες (κακοήθη εξιδρώματα, μη κακοήθη εξιδρώματα και διϊδρώματα) βρέθηκε μαθηματικό μοντέλο το οποίο μπορεί να κατατάξει περίπου το 70% των νέων περιπτώσεων χρησιμοποιώντας μόνο πέντε πρωτεΐνες του ασκιτικού υγρού: την ολική πρωτεΐνη, την LDH, τον TNFa, το C4 and τις απτοσφαιρίνες.

Σημείωση

Σ.Κ. = Σχετικός Κίνδυνος

Δ.Ι.= Διάστημα Εμπιστοσύνης

FIGURES

TABLES

1. INTRODUCTION

1.1. Cirrhosis aetiologies

Cirrhosis of the liver is the leading nonmalignant cause of death amongst digestive diseases throughout most of the developed world (Everhart & Hoofnagle, 1992). The term 'cirrhosis' is a pathological term, being in itself not really a single disease but a consequence of several diseases that differ greatly in their pathogenesis, natural history and response to treatment. Cirrhosis has been defined by a World Health Organisation (WHO) group as "a diffuse process characterized by fibrosis and the conversion of the normal liver architecture into structurally abnormal nodules" (Anthony et al, 1977, Anthony et al, 1978 cited on pages 397 & 323-4, MacSween et al, 1994). Cirrhosis is often classed by its aetiology: drugs and toxins (e.g. alcohol), infection (e.g hepatitis B and C), autoimmune (e.g. autoimmune hepatitis, primary biliary cirrhosis), metabolic (e.g. Wilson's disease), biliary obstruction, vascular or cryptogenic. Cirrhosis is classified as being of cryptogenic aetiology if there is no recognizable cause for the disease.

In the 35 to 54 year age group in the USA, cirrhosis, predominantly alcoholic, was found to be the fourth most common cause of death in males and the fifth most common in females (Galambos et al, 1985, cited on pg 323, MacSween et al, 1994). Alcohol intake has been found to be the main determinant of liver cirrhosis in Italy and other Western countries (Corrao and Arico, 1998). Only 10% to 25% of alcoholics, however, suffer from liver cirrhosis during their lifetime. A prospective study of 258 men with an average daily alcohol intake of over 50g for more than one year, with a follow-up of between 10 and 13 years, found a development of cirrhosis of approximately 2% per year (Sørensen et al, 1984). No relationship was found between the average daily alcohol consumption among abusers and the rate of subsequent development of cirrhosis, but intermittent abusers had a lower rate of cirrhosis than daily abusers. It has been found that alcohol and the hepatitis C virus (HCV) interact synergistically on the risk of liver cirrhosis, particularly when high levels of alcohol intake are present (Corrao and Arico, 1998). Primary liver carcinoma has been estimated to develop in between 5% and 15% of patients with alcoholic cirrhosis (F.I. Lee, 1966, Hislop et al, 1982, cited on pg 327, MacSween et al, 1994). In the present study, male cirrhotics were diagnosed as having cirrhosis due

to alcohol abuse if they stated having had an average daily alcohol intake greater than 40g for more than 5 years, without any other causative factor for cirrhosis. The criteria were similar for female cirrhotics but with an upper limit of 20g instead of 40g.

Hepatitis B virus (HBV) infection is responsible for at least 30% of non-alcoholic cirrhosis, and chronic liver disease due to HBV is one of the main causes of death worldwide (Realdi et al, 1994). Estimates on the rate of progression of hepatitis B are, however, contradictory (Di Marco et al, 1999). There are about 400 million chronic carriers of HBV throughout the world (Bergsland & Venook, 2000). Although chronic hepatitis B virus carriage is common, only about 5% of adult subjects who become infected with HBV develop chronic infection (Bergsland & Venook, 2000). The chronic infection rate for infected neonates is much higher. The highest rates of chronic carriage are in Southeast Asia and sub-Saharan Africa where 8 to 10 percent of the population become chronically infected (WHO, fact sheet 204, 2000). High rates of chronic carriage are also found in southern parts of Eastern and Central Europe (WHO, fact sheet 204, 2000). A vaccine which is 95% effective in preventing the development of chronic infection has been available since 1982 and at present forms part of the immunisation programme of at least 116 countries (WHO, fact sheet 204, 2000), Greece included.

The usual endpoints for successful treatment of chronic hepatitis B are not prolonged survival or prevention of cirrhosis but rather a decline in serum aminotransferase activities and a disappearance of viral markers such as HBV DNA and hepatitis B e antigen (HBeAg) (Everhart & Hoofnagle, 1992). The persistence of HBV DNA in serum measurements has been found to be strongly correlated with the development of cirrhosis (Fattovich et al, 1991 cited in Everhart & Hoofnagle, 1992). It should be noted that termination of HBV-replication implies that serum HBV DNA becomes negative, although this does not necessarily imply that the patient is no longer infected with HBV as the virus may persist elsewhere e.g. in pancreatic cells or brain cells.

The WHO estimates that 170 million people (3% of the world's population) are infected with HCV (WHO, 1999, cited in Wasley and Alter, 2000). On the basis of studies using blood donors, the lowest anti-HCV prevalence rates (0.01 to 0.1%) have

been reported from the UK and Scandinavia whilst intermediate rates (1 to 5%) have been reported from Eastern Europe, the Mediterranean, the Middle East and regions of Africa and Asia (Wasley and Alter, 2000). By far the highest rates have been reported in Egypt, estimates being between 17% and 26% (Wasley and Alter, 2000). It should be noted that estimates based on blood donor populations are underestimates of true infection rates e.g. in the US the prevalence of HCV infection amongst volunteer blood donors in 1990 was 0.6% whilst in the general population the prevalence rate was estimated be 1.8% (Wasley and Alter, 2000). On Crete, the prevalence of HBsAg is lower than in mainland Greece (and more closely resembles the situation in Spain and Japan) whilst anti-HCV positivity has been reported to be higher in the general population in Crete than in mainland Greece (Lionis et al, 1997a, Lionis et al, 1997b, Fragiadakis et al, 1996).

Approximately 80% of patients infected with HCV develop chronic infection (WHO, fact sheet 164, 2000). This difference in outcome between HCV and HBV is thought to be determined by variations in the cell-mediated immune response (Bergsland & Venook, 2000). In contrast to HBV, no vaccine is currently available to prevent hepatitis C (WHO, fact sheet 164, 2000). It has been reported that the majority of chronically infected subjects do not develop symptomatic liver disease from HCV for at least thirty years after infection, if at all (Berk, 2000). WHO estimates for the development of cirrhosis in subjects with chronic infection are 10% to 20% over a thirty-year period (WHO, fact sheet 164, 2000). A study of patients with chronic hepatitis C, however, reported a risk of cirrhosis of 40% at 5 years and 60% at 8 years with an average annual rate of 8% (Tremolda et al, 1992, cited in Fattovich et al, 1997). Between 1% and 5% of persons with chronic infection develop liver cancer over a twenty to thirty year period (WHO, fact sheet 164, 2000).

In the present study, cirrhosis patients were classed as having hepatitis B as cirrhosis aetiology if they were found positive for the hepatitis B surface antigen, HBsAg. These cases are also referred to as cases of **cirrhosis type B**. Hepatitis B markers were detected with commercial (Abbott Laboratories) enzyme linked immunosorbent assay (ELISA) kits. The aetiology of cirrhosis was taken to be the HCV if the patient was found to be positive for the presence of HCV antibody (anti-HCV) in serum. These cases are also referred to as cases of **cirrhosis type C**. The diagnostic test used

for the detection of anti-HCV was ELISA. In the epidemiological surveys undertaken to determine the prevalence of viral markers, serum samples were tested were tested by microparticle capture enzyme immunoassay using Abbott kits (North Chicago, Il): ImxHBsAg (hepatitis B surface), IMxCORE for total HBCAb and ELISA 2 and ELISA 3 for anti-HCV.

1.2 Primary biliary cirrhosis (PBC)

Primary biliary cirrhosis (PBC) is a chronic cholestatic progressive liver disease of unknown aetiology, predominantly affecting middle-aged women. Cirrhosis lesions are found only in stage IV of the disease whilst in the previous three stages there are lesions in the small intrahepatic bile ducts, leading to their progressive elimination. The diagnosis of PBC is most often made when the patient is still asymptomatic, with abnormal liver biochemistry and/or antimitochondrial antibodies (AMA) detected in the blood at the time of a routine check-up or following blood tests for a related disorder (Heathcote, 2000). PBC is typically characterized by the presence of AMA in the serum. Several investigators have reported patients, however, who clinically, biochemically and histologically have all the features of PBC but whose sera consistently tests negative for AMA (Heathcote, 2000). These patients have been described as having **autoimmune cholangitis** (**AIC**). They are most likely to be cases of PBC except that their non-organ-specific antibody profile is more like that found in autoimmune hepatitis (Heathcote, 2000).

The age- and sex-adjusted prevalence of PBC in a U.S. community (Olmsted County) in 1995 was 40 cases per 100,000 population with a 95% confidence interval (C.I.) of 27 to 53 cases per 100,000 population; prevalence rates were greater among women than men, with rates of 65 and 12 per 100,000 respectively (Kim et al, 2000). In European studies, PBC prevalence rates have been estimated to be between 0.5 and 39 per 100,000 population (Kim et al, 2000). The lower European rates may be due to an underlying genetic influence (supported by the clustering of cases within families).

It was seen twenty years ago that the hydrophilic bile acid ursodeoxycholic acid (UDCA) improved the serum biochemistry of patients with autoimmune hepatitis (Heathcote, 2001). It is thought that in PBC, long-term treatment with UDCA might displace endogenous bile acids from the enterohepatic circulation and thus reverse

their suspected toxicity (Poupon et al, 1997). Recent guidelines provided by the Practice Guidelines Committee of the American Association for the Study of Liver Diseases (AASLD) recommend that 'appropriately selected patients with PBC with abnormal liver biochemistry should be advised to take UDCA, 13 to 15 mg/kg daily' (Heathcote, 2000). This treatment is recommended both for symptomatic and asymptomatic patients (Heathcote, 2001).

There have been twelve published randomized controlled trials of UDCA therapy in PBC patients to date. A combined analysis of three of the largest trials with 273 UDCA-treated patients and 275 patients administered a placebo (combining Heathcote et al, 1994, Poupon et al, 1994 and Lindor et al, 1994) found that survival free of liver transplantation was significantly improved in UDCA-treated patients as compared to patients originally assigned to placebo (following up to 4 years after the start of treatment) with a reported relative risk of 1.9; 95% C.I. 1.3 to 2.8 (Poupon et al, 1997). A double-blind placebo controlled trial undertaken by the UDCA-Cooperative Group from the Spanish Association for the Study of the Liver assessing the long-term effects of UDCA with 192 patients (99 UDCA, 93 placebo) did not detect a significant difference in the times to death or liver transplantation between the two groups (Parés et al, 2000). A meta-analysis of 11 of the randomized controlled trials found no evidence of a therapeutic benefit of UDCA in PBC (Goulis et al, 1999). For the meta-analysis, however, evaluation was only of treatment given for between six months and two years. In the absence of a control group, as with Cretan PBC patients who are all currently offered UDCA treatment, survival in PBC patients is often compared with survival predicted by the Mayo model (for further details of the Mayo model see **Section 2.1.5.**). The option of a liver transplantation is not yet available in Crete.

1.3 Decompensation in cirrhosis patients

Decompensated liver disease for the purposes of the present study of cirrhotic patients was defined, following Bonis et al (1999), as the presence of at least one episode of ascites, jaundice, hepatic encephalopathy or gastrointestinal bleeding of variceal origin. Each of these factors is weighted equally. **Ascites** is the development of fluid within the abdominal cavity. Ascites is a common clinical finding caused mainly by advanced liver disease or malignant neoplasms in the abdominal cavity.

The abdominal cavity fluid may be a transudate (protein concentration <30 g/L) or exudate (protein concentration >30g/L) (pg 83, Hayes & Simpson, 1995). Although estimation of the albumin concentration of the ascitic fluid provides an easy method for separation of transudates from exudates, a practical solution for the separation of malignant from non-malignant exudates does not yet exist (Alexandrakis et al, 2000). A variety of laboratory tests have been evaluated for their ability to improve the accuracy of differential diagnosis between benign and malignant ascites, but complete discrimination has not been achieved (Bansal et al, 1998, Gupta et al, 1995, Gerbes et al, 1991, Jungst et al, 1986, Scholmerich et al, 1984). Cytological examination of ascitic fluid, despite its high specificity may produce considerable false -negative results, with sensitivity ranging from 40% to 60% (Garrison et al, 1986 and Dekker et al, 1978 cited in Alexandrakis et al, 2001). A significant discriminatory efficiency has been ascribed to the serum ascites albumin concentration gradient (Lee et al, 1992), ascitic fluid lipids (Jungst, 1986) and ascitic fibronectin (Lee et al, 1992) but these results have not yet been confirmed by other studies (Mauer, 1988 cited in Alexandrakis et al, 2001). Serum levels of vascular endothelial growth factor (EGF) can be used to monitor the clinical course of ovarian cancer patients, but cannot be used for discriminatory purposes (Yamamoto et al, 1997).

The biochemical analyses undertaken for the ascites patients in the present study involved the collection of ascitic fluid in sterile ethylene diamine tetra acetic- (EDTA-K_3) vacutainer plastic tubes (Becton Dickenson, NJ USA) for cell counting. The tubes were centrifuged at 3000 rpm for 10 min at 0^oC to obtain cell-free supernatants that were subsequently stored at -70^oC until assayed. Blood sera were obtained from peripheral venous blood by allowing the blood to clot at room temperature for between one and two hours. These samples were centrifuged at 3000 rpm for 20min. Serum samples were stored at -80^oC until assayed. Routine biochemical parameters were measured in the serum and the ascitic fluid using a RA-1000 autoanalyzer (Technicon Instruments Corporation, New York, U.S.A). Total protein concentration was determined by the Biuret reaction. Albumin concentration was measured with the bromocresol green method. Ferritin and cytokines IL-1a, IL-1b, IL-2, IL-6 and TNF-a were assessed by immunoradiometric assay kits (Amersham Int, UK) IL-8 was measured by an enzyme-linked immunosorbent assay (ELISA) kit (Amersham Int,UK). The immunoglobulins IgG, IgA, IgM and the acute phase proteins HAP,

TRF, a_1AT, a_1AG, CER, CRP and a_2MG were determined by nephelometry (Kallestad, model QM300, MN USA). Complement factors C3 and C4 were measured by radial immunodiffusion (Biomerieux 69280 MARCY L' Etoile/France). Futher details of all biochemical analyses undertaken for the ascites patients in the present study (**Section 3.4**) can be found in our publications: Alexandrakis et al (2000) and Alexandrakis et al (2001).

1.4 Hepatocellular carcinoma

Primary hepatocellular carcinoma (HCC) is a frequent complication of chronic liver disease and is the most common form of primary liver malignancy. It is estimated to be the seventh most common cancer in men and the ninth in women worldwide (El-Refaie et al, 1996). There is, however, significant geographical variation in the prevalence of HCC. In Europe and North America the prevalence is estimated as 2 to 4 per 100 000 population whereas in some regions of Asia and Africa it is more than 100 per 100 000 population. In Greece, the average annual mortality rate from primary HCC has been estimated, based on death certificates from 1971 to 1973, to be 23.3 in males and 14.0 in females per 100,000 (Trichopoulos et al, 1975). More recent data from Crete provide crude and age-standardized liver cancer mortality rates of 11.44 and 9.76 per 100,000 in males and 8.18 and 5.44 per 100,000 in females respectively (1994 data, Cancer Registry of Crete, www.med.uoc.gr/~biostats/gastroint1.htm).

An HCC incidence rate varying between 3% and 6.5% per year of follow-up has been reported in cirrhotic patients (Bolondi et al, 2001, Fasani et al, 1999). Cirrhosis aetiology is one of the major risk-determining factors for HCC development: HCC is rare in patients with PBC, of higher frequency in alcoholic cirrhosis and of higher still frequency in cirrhosis due to chronic HBV infection (pg 419, MacSween et al, 1994). Over 70% of HCC patients in Western countries have underlying liver cirrhosis (Badvie, 2000). In areas with an intermediate incidence of HCC (5 to 20 cases per 100,000 individuals), the neoplasm is associated with liver cirrhosis in more than 90% of cases (Bolondi et al, 2001). Crude and age-standardized liver cancer incidence rates in Crete have been estimated as 15.5 and 13.6 per 100,000 in males and 9.7 and 6.9 per 100,000 in females respectively (1994 data, Cancer Registry of Crete, www.med.uoc.gr/~biostats/gastroint1.htm).

The most important HCC predisposing factors are liver cirrhosis and chronic hepatitis B or C viral infection, with infection with the hepatitis B virus being reported to be the cause of 80% of HCC cases worldwide (Badvie, 2000). Sequences of the HBV genome have been found to be integrated at the DNA of both malignant and normal hepatocytes despite the absence of serological markers (Diamantis et al, 1992, Paterini et al, 1993). Recently, the presence of HCV has been found in certain countries to be as important a risk factor for HCC as HBV (Chiba et al, 1996), with HCV genotype 1b being particularly associated with HCC (Tanaka et al, 1996). Reports from Japan, Spain and Italy, countries which have an intermediate prevalence of HBV, show anti-HCV to be present in high proportions of HCC patients, ranging from 65% association of HCC cases with HCV in Italy to 75% association in Spain (Kiyosawa, et al 1990, Bruix et al, 1989, Colombo et al, 1989). In the U.S., persistent hepatitis C virus infection is the cause of 30% to 50% of cases of HCC (Bergsland & Venook, 2000). In countries where HBV is highly endemic, such as Korea and South Africa, the association of HCC with HCV is relatively low, ranging between 17% and 29% of the total (Lee et al, 1993, Kew et al, 1990).

On mainland Greece, HCC has been found in previous studies to be mostly associated with HBV (Trichopoulos et al, 1978, S Hadziyiannis, 1980). A more recent study by Manesis et al (1995) found 62% association with HBV and only 13% of cases being anti-HCV positive. The interpretation of these associations, however, is somewhat limited by the fact that the study was a randomized controlled trial of 85 HCC patients with advanced inoperable disease. In Crete, HCC is thought to be mostly associated with HCV, unlike mainland Greece, with a 54% association found in our (published) study of the natural history of HCC in Crete (Kouroumalis et al, 1997).

Potentially curative therapeutic options for HCC patients include complete surgical resection, which is performed mainly in patients with well-preserved liver function, and orthotopic liver transplantation in patients with advanced liver impairment (Michel S et al, 1997, Philosophe B et al 1998). In most patients, however, the tumours at presentation are inoperable and the prognosis is poor (Kouroumalis et al, 1997). Treatment of such hepatocellular carcinomas has, in general, been unsatisfactory (Simoneti et al, 1997, Raoul et al, 1999, Liu al, 2000, Liovet et al,

2000). Other treatment options include systematic chemotherapy, targeted (tumour-specific) and rational biologic therapies, cryotherapy, immunotherapy and hormonal therapy (Badvie, 2000). The use of these options is limited, however, by the current lack of definitive data on their efficacy (Badvie, 2000, Bergsland & Venook, 2000). **Somatostatin** is a hormone with known antimitotic activity in various neoplasms (Kouroumalis et al, 1998). Its synthetic analogues have been shown to delay tumour growth in animals (Schally, 1988) and it is thought that they may temporarily inhibit tumour growth in humans (Wood, 1996). In fact, the rationale behind testing the effect of somatostatin analogues in patients with HCC, be this is in the form of short- or long-acting analogues, is that it is likely that its actions on a molecular level eventually lead to a shrinkage of human tumour cells (D Shouval, 1998). The first somatostatin analogue introduced for clinical use was **octreotide** (Wood, 1996). Another cyclic analogue with a slightly different activity profile is **lanreotide** (Wood, 1996).

Diagnosis of HCC at the University Hospital is based on either histology or an alpha-fetoprotein concentration greater than 500 ng/ml plus compatible liver imaging and/or selective angiography (as reported in Kouroumalis et al, 1997). A radioimmunoassay method was used for AFP estimation in the HCC octreotide trial patients. All percutaneous liver biopsies were ultrasound-guided.

1.5 Survival analysis: terminology and single sample inference

Medical studies often involve data on occurrences of a point event such as death. These data can be analysed by logistic regression or other binary variable techniques, but these methods do not make full use of the available information. In survival analysis, the time to occurrence of an event (or 'time to failure'), as well as the fact that the event happened, is assessed. One question which may be asked is: why are survival data not amenable to standard statistical procedures? The main distinguishing feature of survival data is the presence of **censoring**. A subject is considered censored if, for any reason, a survival time (known more generally as 'failure time') is not observed . There are many different types of censoring. The following terms are commonly seen in the literature:

1) **Right censoring** (which is the most common type of censoring and the type observed in the data to be analysed in the present study): censoring occurs after an

individual has been entered into a study i.e. to the right of the last known survival time. The right-censored survival time is less than the true, but unknown survival time. Two types of events that may lead to right-censored observations are the end of the study and loss-to-follow-up e.g. it is known that a subject was alive 3 years ago but there is no information after that date.

2) **Informative censoring**: here, the prognosis for an individual who is censored at time **c** is not the same as that of any individual who has survived to **c** e.g. the survival time is censored because treatment is withdrawn as a result of deterioration of the physical condition of the patient. See **Section 2.1.2.2.3.** for details of how to test for the presence of informative censoring.
3) **Type I and Type II censoring**: these are types of right censoring. It is assumed that the trial has **n** individuals at the time origin, t=0. **Type I** censoring occurs when the experiment is stopped at a fixed time T_{max}. i.e. censoring time is fixed in advance. The data set then consists of 2 groups: m individuals in whom failure has occurred, and n-m who have not yet failed. The total number of observed failures m is a random variable (r.v.). In **Type II** censoring the experiment stops after the event has occurred in a fixed number (k) of subjects. In this case, $T_{max} = T_k$ and the failure time of the k^{th} subject is random. **Type II** censoring is an example of a general scheme known as **evolutionary censoring**, in which censoring may depend on the past but not on the future of the process (pg 5, Cox & Oakes, 1984). **Type I** and **Type II** censored observations are also known as **singly censored data**.
4) **Type III censoring, or random censoring**: this is also a type of right censoring. It occurs when the period of the study is fixed and patients enter the study at different times during that period. **Type III** censored observations are also known as **progressively censored data**.

Let $\mathbf{T_1, T_2, T_3,\ldots,T_n}$ be the independent, identically distributed (iid) survival times for the n individuals in the study, if they were all under observation till they died. Consider right censoring in which the individual leaves the trial at time c_i and either the survival time T_i is known, if $T_i \le c_i$,or it is known that $T_i > c_i$. Let $\mathbf{c_1, c_2, c_3,\ldots, c_n}$ be the corresponding iid censoring times if no subjects died (i.e. observation ceases at this time). So the observations consist of $\mathbf{Y_i=min\ (T_i,c_i)}$ together with the indicator

variable δ_i=**1 if** $T_i \leq c_i$ (uncensored) and δ_i=**0 if** $T_i > c_i$ (censored). **Random censoring** assumes T_i and c_i are independent random variables i.e. censoring is 'uninformative' as knowing the distribution of $\mathbf{c_i}$'s provides no information on the survival times **T** (see definition **2)** above). This definition includes **Type I** censoring (but not **Type II** censoring). In some cases, all the $\mathbf{c_i}$'s may be known. In **Type I** censoring, all the $\mathbf{c_i}$'s are equal i.e. $\mathbf{c_i} = \mathbf{c}\ \forall \mathbf{i}$. A vital assumption for **random censoring** is that, conditional on the explanatory variables, the prognosis for an individual who is censored at time **c** is the same as that of any individual who has survived to **c**. Most analyses (including all those based only on likelihoods) are valid under a weaker assumption: that of **independent censoring**, in which the hazard (risk of death) at time **t** conditional on the entire history depends only on the survival of that individual to time **t** (pg 268, Venables & Ripley, 1994). This includes **Type II** censoring.

The time to failure for a single individual can be represented mathematically by a non-negative r.v. **T**, which measures the length of the interval between a point of origin and an end point. In the present context, the three end-points considered are death, the occurrence of decompensation and the occurrence of HCC. T is measured in months in all three cases. The **survivor function S(t)** is defined as $S(t) = P(T \geq t)$ and is the marginal probability of being event-free up to time t. S(t) can be thought of as the proportion of individuals still alive at time t. The **cumulative distribution function** F(t) of T is $F(t) = 1 - S(t) = P(T < t)$ and the **probability density function** (p.d.f.) is f(t)=dF/dt if T is continuous (or f(t)=P(T=t) if T is discrete). The **hazard function h(t)** (also known as the instantaneous failure rate or age-specific failure rate) describes the instantaneous risk of failure at every time t, given that failure has not occurred prior to that time i.e. it is a limiting probability of death at time t, conditional on survival to that time, $h(t) = \lim_{\delta t \to 0} \left\{ \frac{P(t \leq T < t + \delta t / T \geq t)}{\delta t} \right\}$. h(t) represents the **death rate at time t**, the rate at which deaths occur divided by the proportion of the population still surviving, so $h(t) = \frac{f(t)}{S(t)}$. It should be noted that the hazard function is not a probability but a death rate per unit of time so it is not necessarily less than 1. In continuous time, it can be shown by integration that $S(t) = e^{-H(t)}$ where

$H(t) = \int_0^t h(u)du$. H(t) is called the **cumulative hazard function** or **integrated hazard function**. S(t) and h(t) are usually estimated from the observed survival times. The hazard function plays a central role in survival analysis in that the precise nature of f(t) or S(t) is often not known but there is information on how the failure rate will change over time. In the investigations comprising the present thesis, h(t) measures the "risk" of dying (for cirrhosis patients), of decompensating (for compensated cirrhotics) and of HCC occurring (cirrhosis patients).

The exponential distribution is the simplest distribution that can be applied to survival data (D. Oakes, pg 111, 1991). It applies when the lack-of-memory property holds: P(T>t+s / T>s)=P(T>t) i.e. the conditional distribution of the remaining time to failure T-s, given survival to time s, does not depend on the time point s. Under the exponential distribution, the hazard function is therefore constant, h(t)=λ and $S(t) = e^{-\lambda t}$. A distribution that is not as restrictive as the exponential is the Weibull distribution with index λ>0 and scale parameter ρ>0: $S(t) = e^{-(\rho t)^{\lambda}}$, $h(t) = \lambda\rho(\rho t)^{\lambda-1}$. This distribution includes the exponential as a special case (λ=1) and has a monotone hazard, increasing if λ>1 and decreasing if λ<1. A Weibull distribution is commonly used in carcinogenesis models for time to incidence (as multi-stage theories suggest a power law for the hazard function) (D. Oakes, pg 112, 1991). Commonly used estimation methods, including those used in the present thesis, do not require knowledge of the p.d.f. of T; they are known **as non-parametric methods**. These are described in detail in **Section 2.1.1.**

For a continuous survival distribution, a subject observed to fail at time t contributes a term f(t) to the likelihood (the density of failure at t) and a subject censored at time c contributes S(c). The likelihood for a sample of size n (n independent subjects indexed by i) under independent right censoring, when T has density f(t;θ) depending on the parameter θ, can be written as $lik = \prod_u f(t_i;\theta)\prod_c S(c_i;\theta)$, with u representing all uncensored and c all censored subjects (the definition of the simple **likelihood** being the joint density of the observed values considered as a function of the unknown parameters). In terms of the observed time y_i=min (t_i,c_i) the log likelihood is

$l = \sum_u \log f(y_i;\theta) + \sum_c \log S(y_i;\theta)$ which, using h(t)=f(t)/S(t), may be written as $l = \sum_u \log h(y_i;\theta) + \sum_{all} \log S(y_i;\theta)$ (using the identity log(AB)=log(A)+log(B)). Maximum likelihood estimates of the parameters can be found by maximizing the log likelihood. For a single homogeneous random sample, the hypotheses with regard to the parameters of the survival time distribution are classically tested with the likelihood ratio, Wald or score statistics (details of these three types of tests are given in **Section 2.1.2.1**).

1.6 Statistical software

In the present thesis, the statistical analyses were undertaken on a p.c. using a combination of two software packages: S-PLUS version 4 (MathSoft Inc.) and the Statistical Package for the Social Sciences, SPSS version 8.0 (SPSS Inc.) on a P.C. S-PLUS is based on the S language and has more extensive statistical capabilities than SPSS. Interactive programming using S functions will be used in the present study with details provided in the **Appendix**. The software package most commonly used by medical researchers at the University of Crete for statistical analyses seems to be SPSS. The **Methods** section contains descriptions of SPSS functions and capabilities, with specifics provided in the **Appendix**. Survival analyses were undertaken using either SPSS or S-PLUS. Prognostic indices were constructed using SPSS and Microsoft Excel. S-PLUS was used for bootstrapping, jackknifing and for the construction of tree models. SPSS was used for discriminant analysis.

1.7 Study aims

The main aims of the series of investigations comprising the present thesis are to develop and validate prognostic models for cirrhosis and HCC patients using survival analysis and discrimination techniques. In general, prognostic models are used to investigate patient outcome in relation to patient and disease characteristics (Altman & Royston, 2000). The development of prognostic classification systems for use as clinical prediction rules is of major interest in many areas of clinical research (Graf et al, 1999). There are two main ways in which a prognostic model may be useful:

a) it may be used to estimate the prognosis of patient groups and individual patients in terms of their survival times

b) it may allow the reliable classification of patients into two or more groups with different prognoses. These classification methods can be used to avoid unnecessary tests and to influence the type of therapy offered to patients.

The rates of development of decompensation and HCC in cirrhosis patients in general and in Cretan cirrhotics in particular, and the degree to which they contribute to the mortality rate, are relatively unknown. In the present study, the models developed will include predictions and comparisons of times to decompensation and survival times for compensated and decompensated cirrhotics and time to HCC incidence models for cirrhosis patients. In addition, survival models will be developed for HCC patients and discrimination models will be developed for patients with ascites.

The present study entails data on all patients with cirrhosis presenting at the Gastroenterology Clinic after its opening and before the end of the year 2000. The patient groups studied were the following:

a) patients diagnosed at the Clinic with cirrhosis: the natural history of cirrhosis in Crete, according to disease aetiology.

b) Type C cirrhosis patients: treatment with ***plaquenil*** versus no treatment.

c) PBC patients treated with ***UDCA***: comparisons with the widely-applied prognostic Mayo model

d) patients in whom HCC has developed: the natural history of HCC

e) HCC patients treated with the ***short-acting somatostatin analogue octreotide***: a randomised, controlled study of treated HCC patients versus untreated patients and

f) HCC patients treated with a ***long-acting somatostatin analogue***: a study of treated HCC patients compared with historical controls.

g) patients who have developed ascites: i) discrimination between malignant and non-malignant cirrhotic ascites ii) discrimination between 3 types of peritoneal effusions in ascites patients.

There are four specific questions of interest with regard to the natural history of cirrhosis (patient group **a)**):

1) What is the average time until decompensation for patients presenting with compensated cirrhosis, and how is this influenced by cirrhosis aetiology and other prognostic factors?

2) What is the average survival time for patients presenting with compensated cirrhosis, and to what extent is this influenced by cirrhosis aetiology, other prognostic factors and whether decompensation occurs?
3) What is the average survival time for patients presenting with decompensated cirrhosis, and how is this influenced by cirrhosis aetiology and other prognostic factors?
4) What is the HCC incidence rate in Cretan patients with cirrhosis?

Chapter 2 contains details of the statistical methods used both for the survival analyses and for the discrimination between diagnostic groups. The data used in the statistical analyses are described in **Chapter 3**, in which details are provided of the patients involved, the design of each of the studies comprising the thesis and the measurements available for each of the patient groups. **Chapter 4** contains the results of all analyses. The results are divided into sections, corresponding to the data sets **a)** to **g)** above. In addition, results are presented from three epidemiological surveys undertaken into the prevalence of HBV and HCV viral markers in Crete. An interpretation of the results, comparisons with other studies and presentation of possible study limitations are provided in **Chapter 5**.

2. METHODS

2.1 Survival analysis

2.1.1 Non-parametric estimation methods and their application

2.1.1.1.The Kaplan-Meier product-limit method

The most widely used non-parametric method for estimating the survivor function is the Kaplan-Meier product-limit (PL) estimator of survival (Kaplan & Meier, 1958). The Kaplan-Meier PL estimate can be derived simply with the use of conditional probabilities for the set of **n** individuals with **s** (**s ≤ n**) observed failure times $\mathbf{t_{(1)} < t_{(2)} <\ldots< t_{(s)}}$, using $S(t) = P[T \geq t] = P[T > t_{(1)}] \times P[T > t_{(2)} \,|\, T > t_{(1)}] \times \ldots \times P[T > t_{(s)} \,|\, T > t_{(s-1)}]$ where $t_{(s)}$ is the largest observed failure time prior to time t. Each conditional probability can be thought of as $\frac{\text{number alive just after } t_{(j)}}{\text{number alive just before } t_{(j)}} = \frac{r_j - d_j}{r_j}$ where r_j represents the number of individuals at risk of failure just before time $t_{(j)}$. By convention, if any subjects are censored at time $t_{(j)}$ then they are considered to have survived for slightly longer than the deaths at $t_{(j)}$. The **Kaplan-Meier product-limit estimator of survival** is therefore

$\widehat{S}(t) = \prod_{j=1}^{s} (1 - \frac{d_j}{r_j})$ for $t_{(s)} < t \leq t_{(s+1)}$.

As $\hat{S}(t)$ is subject to sampling error, the variance of the estimate can be estimated using **Greenwood's formula** (Greenwood, 1926), as used in SPSS: $\hat{V}[\hat{S}_{KM}(t)] = [\hat{S}_{KM}(t)]^2 \sum_{j=1}^{s} \frac{d_j}{r_j (r_j - d_j)}$ for $t_{(s)} < t \leq t_{(s+1)}$. Single time point survival estimates (e.g. survival after 5 years) can be made using Greenwood's formula to estimate the standard error and from this to obtain C.I.s. To compare entire survival curves, however, a different procedure is needed (see **Section 2.1.1.4** below).

If no censoring is present, the Kaplan-Meier PL estimate reduces to the **empirical survivor function**

$\widetilde{S}(t) = \frac{\text{Number of subjects with survival time} \geq t}{\text{Total number of subjects in the data set}}$ (pg 15, Collett, 1994).

One question sometimes asked is: what is the difference between the PL estimator and the **actuarial estimator** for S(t)? The actuarial estimator is sometimes used to

estimate S(t) from grouped data. In this procedure, failures and censored observations are grouped into a small number of time intervals and the estimated hazards d_j / r_j used in the product limit estimator are replaced by $d_j /(r_j - c_j / 2)$ where c_j is the number of observations censored in the j^{th} interval (Oakes, pg 115, 1991). Essentially, the only difference is that the PL estimate is based on individual survival times whereas the actuarial survival times are grouped into intervals. There are not usually major differences between the two estimates, except if the data are heavily tied (pg 56 Cox & Oakes, 1984). Also, a relatively high risk of early failure may be obscured by the actuarial estimator. A second common question is

"What is the difference between **life-table estimates** and the PL estimates?"

The term 'life-table estimate' is actually synonymous to 'actuarial estimate' (Collett, pg 17, 1994). Basically, the PL estimate can be considered as a special case of the life-table (or actuarial) estimate where each interval contains only one observation (pg 66, Lee, 1992). The life-table/actuarial method requires a fairly large number of observations so that the survival times can be grouped into intervals (pg 86, Lee, 1992). In the present series of investigations, detailed time data are available, rather than only grouped data. Therefore, the Kaplan-Meier PL estimator of survival will be used throughout the present series of investigations, to obtain estimates of survival distributions for groups of cirrhosis and HCC patients.

2.1.1.2 Estimating median follow-up

The most obvious way to calculate the average follow-up time of the patients in a study is to use the median follow-up time of all patients. This is, however, of questionable value as it is directly affected by the times of the observed events (Altman et al, 1995). Presumably, it therefore does not provide an accurate reflection of the average length of time in the study for each patient. To use the median follow-up time of survivors only may be inappropriate as the estimate is unstable if the number of survivors is small. Two more acceptable alternatives when there is a relatively high degree of censoring are, according to Altman et al (1995), to use either the time interval from the median patient entry to the cut-off date of the study or the median time to censoring using a **'reverse' Kaplan-Meier analysis**, exchanging the outcomes 'dead' and 'censored' and taking the 50% point of the resulting curve, as described by Shuster (1991).

It should be noted, however, that the median follow-up is a single measure of follow-up and as such it can only play a limited role, as there are many factors influencing survival curves (Shuster, 1991). When performing survival analyses, the Kaplan-Meier survival curve adjusts for variable lengths of follow-up and provides an unbiased estimate of the true target population survival curve, rendering calculation of follow-up time relatively unimportant. As summaries of follow-up are useful in comparing lengths of different studies, however, they are provided here for each data set analysis. The 'reverse' Kaplan-Meier process was used, unless otherwise stated, given that many of the data sets involved a high degree of censoring.

2.1.1.3 Estimating median survival

The **median survival time** is the time beyond which 50% of the individuals in the population under study are expected to survive (pg 31, Collett, 1994). As the non-parametric estimates of S(t) are step-functions, there will not usually be a realised survival that makes the survivor function precisely 0.5 so the estimated median survival time is defined to be the smallest observed survival (death) time for which the value of the estimated survivor function is less than 0.5. A similar procedure holds for the other percentiles; the p^{th} percentile of the distribution of survival times is defined as t(p) such that F{t(p)}=p/100. Using the estimated survivor function, the estimated p^{th} percentile is the smallest observed survival time $\hat{t}(p)$ such that $\hat{S}\{\hat{t}(p)\} < 1-(p/100)$. Collett (1994) presents the following method for calculating approximate confidence intervals for the percentiles: the standard error (se) of the estimated p^{th} percentile is $se\{\hat{t}(p)\} = \frac{1}{\hat{f}\{\hat{t}(p)\}} se[\hat{S}\{\hat{t}(p)\}]$ where $se[\hat{S}\{\hat{t}(p)\}]$ can be found using Greenwood's formula and an estimate of the p.d.f. at $\hat{t}(p)$ is $\hat{f}\{\hat{t}(p)\} = \frac{\hat{S}\{\hat{u}(p)\} - \hat{S}\{\hat{l}(p)\}}{\hat{l}(p) - \hat{u}(p)}$ where $\hat{u}(p) = \max\{t_{(j)} \left| \hat{S}(t_{(j)}) \geq 1 - \frac{p}{100} + \varepsilon \right.\}$ and $\hat{l}(p) = \min\{t_{(j)} \left| \hat{S}(t_{(j)}) \leq 1 - \frac{p}{100} - \varepsilon \right.\}$ for j=1,2,…,r (where there are r death times and n individuals, n ≤ r)and small ε (often taking ε=0.05 is adequate although larger values may be needed). Then the corresponding 100(1-α)% C.I. has limits $\hat{t}(p) \pm z_{\alpha/2} se\{\hat{t}(p)\}$ where $z_{\alpha/2}$ is the upper α/2 point of the standard normal distribution. This interval estimate is only approximate in that the probability that the interval includes the true

percentile is not precisely 1-α (Collett, pg 34, 1994). SPSS appears to use the above methods for calculation of standard errors and confidence intervals for the median survival time: the manual does not provide details but Collett (pg 310, 1994) states that the SPSS output is the same as that provided by the package BMDP. Alternatives with superior properties exist but they are more difficult to compute.

The **expected survival time** for an individual can be taken as the estimated median survival time, derived from $\hat{S}(t)$, where $\hat{S}(t)$ may be obtained using the Kaplan-Meier method described above or from Cox regression procedures as described in **Section 2.1.2.** The methods described in the previous paragraphs make no assumptions about the functional form that the survival distribution would take in the absence of censoring (pg 48, Cox & Oakes, 1984): they are known as **single sample non-parametric methods**.

2.1.1.4 The log rank test

The non-parametric **log rank test** (Mantel & Haenszel, 1959) is the significance test most commonly used to compare two or more groups of survival data without making any assumptions about the shape of the survival curve. With the log rank procedure, the duration of the experiment can be divided into intervals, the width of which may be determined by the occurrence of deaths (or other end-point). For each time interval and each group, the number of deaths and the number of those who leave the interval alive is calculated. If there are two groups, with one undergoing treatment and the other being a control group, under the null hypothesis of no treatment difference (i.e. the risk of death being equal for the two populations), the observed number of deaths in each interval should be divided between the groups in proportion to the number of subjects at risk at the start of the interval (this gives the **expected number of deaths** for that interval). Within each interval, the expected number of deaths can be calculated, compared with the observed number of deaths and summed over all time intervals. The null hypothesis that the risk of death at any time is equal for the two groups is tested using the standard chi-square test. If the deviations are too large to be explained by chance, there is evidence of a treatment difference (Stablein et al, 1981). In mathematical notation, for the two groups, let r_{1i} and r_{2i} be the numbers of patients alive and not censored in groups 1 and 2 just before time t_i with $r_i=r_{1i}+r_{2i}$. Let $d_i=d_{1i}+d_{2i}$ be the number of individuals who die at t_i in the two groups combined. Let

c_{1i} and c_{2i} be the numbers censored in each group in the previous time interval. Then, at the next death time $t_{(i+1)}$, $r_{1(i+1)}=r_{1i}-d_{1i}-c_{1i}$ for group 1 and similarly for group 2. For each t_i, the probability of death for each subject under the null hypothesis is calculated as $p_{di} = \frac{(d_{1i} + d_{2i})}{(r_{1i} + r_{2i})}$. For each group, the expected number of deaths at each t_i is $e_{1i}=p_{di}$ x r_{1i} and $e_{2i}=p_{di}$ x r_{2i}. Then the total expected numbers of deaths E_1 and E_2, assuming an equal risk of dying at each time in both groups is calculated as $E_1 = \sum_i e_{1i} = \sum_i \frac{r_{1i} d_i}{r_i}$ and $E_2 = \sum_i e_{2i} = \sum_i \frac{r_{2i} d_i}{r_i}$ It should be noted that the total observed number of deaths, $O=O_1+O_2=E_1+E_2$, (where $O_1 = \sum_i d_{1i}$ and similarly for O_2) so it is only necessary to calculate E_1. Then the appropriate test is a chi-squared test on 1 df (as there is one constraint, that the two frequencies add to the sum of the expected, so 1 df is lost giving 2-1=1 df). Further details are given in Machin & Gardner (pg 64-65, 1989) and Bland (pg 284-288, 2000). In some cases, the expected number of deaths in a group maybe larger than the number of individuals starting in the group, so a more accurate description than "expected number of deaths" is "the extent of exposure to the risk of death" (Peto et al, 1977 cited in Armitage &Berry, 1987, pg 430)

The **hazard ratio** can be obtained from the calculations as it is the ratio of the observed to the expected number of deaths in the first group divided by the same ratio in the second group (p288, Bland, 2000): $h = \frac{O_1 / E_1}{O_2 / E_2}$

The Kaplan-Meier approach, with the associated log rank test is, however, limited in its ability to fully describe and model a given data set. Therefore, regression models such as the Cox regression model (Cox, 1972) in which adjustments can be made for other prognostic variables, are also considered (see **Section 2.1.2.1.**). The log rank test is equivalent to using the **score test** of the null hypothesis of equal hazards in the Cox regression model (see **Section 2.1.2.1.** for a definition of the score test) (pg 254, Collett, 1994). In fact, the log rank test has been derived as a test of the null hypothesis $\psi=1$ under the proportional hazards model $h_1(t) = \psi h_0(t)$ (see **Section 2.1.2.1.**) (pg 105, Cox & Oakes, 1984). For alternatives in which the hazards are non-

proportional (e.g. $h_1(t)>h_0(t)$ for $t<t_j$ and $h_1(t)<h_0(t)$ for $t>t_j$), its properties may be poor e.g. the test result may be negative even though the survival curves have entirely different shapes. To ensure that this is not the case, one could consider introducing time-dependent variables and/or plotting the cumulative hazards (to see if the proportion difference in hazards is roughly the same at all times). If the mortalities in one group are not a multiple of those in another (i.e. if the alternatives are not within the proportional hazards class), the **generalized Wilcoxon test** is one of the tests considered more appropriate (p124, Cox & Oakes, 1984).

The generalized Wilcoxon test, known as the **Breslow test** in SPSS, is one of the two non-parametric procedures provided in SPSS in addition to the log rank test to assess the null hypothesis of no difference in survivor functions for two groups of data: the Wilcoxon test (mentioned in the previous paragraph). The third test is the **Tarone-Ware test**. All three tests have the general form $U = \sum_{i=1}^{r} w_i (D_i - E_i)$ where D_i represents the observed number of deaths, E_i represents the expected number of deaths and r is the total number of death times in the two groups. The difference between the tests is the weight factor w_i. In the case of the log rank test $w_i = 1$ for all i whereas for the Breslow test, w_i is the number at risk at each time point and for the Tarone-Ware test w_i is the square root of the number at risk. Therefore, the Breslow test gives the most weight to early events (as the number at risk decreases as events occur) and the Tarone-Ware test weights early cases somewhat less heavily than the Breslow test. If the proportional hazards assumption does not hold, the Breslow test may be more powerful than the log rank test but its power is low when the proportion of censored cases is high (Prentice & Marek, 1979, cited in Norušis/SPSS Inc, 1994, pg 281). The log rank test was applied throughout the present study to assess differences in survival distributions for treated versus untreated cirrhosis patients and also to test for possible differences in the survival times of patients with different levels of prognostic factors. e.g. HCC Okuda II patients versus HCC Okuda I patients, HCC males versus HCC females.

2.1.1.5 Estimating survival probabilities

The proportion of subjects **p** surviving beyond any follow-up time t is estimated by the Kaplan-Meier technique as $p = \prod \frac{r_i - d_i}{r_i}$ where r_i is the number at risk just before time t_i (the i^{th} ordered survival time) and d_i denotes the number of deaths at t_i . The estimated standard error (SE) of p is given by $SE = \sqrt{\frac{p(1-p)}{n'}}$ where n' is the **effective sample size** at time t and can be calculated as $n' = \frac{r_i - d_i}{p}$ (Peto et al, 1977) or, alternatively, n'= n-(no. of subjects lost-to-follow-up before time t) (Peto, 1984).

The 100(1-α)% C.I. for the population value of the survival proportion p at time t may be calculated as p-($u_{1-\alpha/2}$ *SE) to p+($u_{1-\alpha/2}$ *SE) where $u_{1-\alpha/2}$ is the appropriate value from the standard normal distribution for the 100(1-α/2) percentile (pg 64, Machin & Gardner, 1989). For example, for a 95% CI, α=0.05 and $u_{1-\alpha/2}$ =1.96. These estimation methods need to be interpreted with caution if n' is less than 10 or p is outside the range 0.1 to 0.9. The times at which survival proportions are to be estimated were chosen using practical conventions: 5-year survival proportions are often quoted in cancer studies. In the present study, percentages of subjects remaining decompensation-free 3-, 5- and 7-years after diagnosis were estimated for compensated cirrhotics (**Table 4.1.1.2**). Also 3-, 5- and 7-year survival percentages were estimated for compensated cirrhotics and for cirrhotics presenting with decompensation (**Tables 4.1.1.6 & 4.1.1.12**). In addition, percentages of cirrhosis patients remaining HCC tumour-free 3-,5- and 7-years after presentation were estimated (**Table 4.1.1.14**). Six- and 12-month survival percentages were estimated for HCC patients treated with ***octreotide*** **(Table 4.1.5.1.1.)** and also those treated with a ***long-acting somatostatin analogue*** **(Table 4.1.5.2.1.)**.

2.1.2 Regression models in survival analysis

2.1.2.1. The Cox Proportional Hazards model

In the Proportional Hazards (PH) model, the hazard of death at time *t* for the i^{th} individual of n subjects in a study can be written

$$h_i(t) = h_0(t)e^{\beta^T X} = h_0(t)e^{\sum_{j=1}^{p}\beta_j x_{ji}} \quad \textbf{(1)}$$

where x_{ji} is the value of the j^{th} explanatory variable X_j, j=1,..,p for the i^{th} individual, $h_i(t)$ is the hazard function of individual i, i=1,…,n and $h_0(t)$ is the baseline hazard function (and gives the hazard when X=0, representing the ageing process of the entire population). When $h_0(t)$ =λ=constant, the above equation defines an exponential regression model (as defined in **Section 1.5**). When no assumptions are made about the form of the baseline hazard function, the PH model is known as a **Cox PH model**, following a paper by D.R. Cox (1972). The Cox PH model is the regression model most commonly applied to survival data (pg 649, MathSoft, 1997) and can be used both in SPSS and S-PLUS for survival analyses.

The PH model is so-called because the hazard ratio $h(t)/h_0(t)$ is constant over time (i.e. the hazards for different sets of covariates remain in the same proportion for all t), as can be seen if **(1)** is written as $\frac{h(t)}{h_0(t)} = e^{\beta^T z}$. For an individual **i**, the effect on $h_i(t)$ of a unit change in the jth covariate, when all other variables are held constant is $\exp(\beta_j)$ and the **relative risk** (or **hazard ratio**) of an individual i with covariate values z_{ji} compared to individual k with values z_{jk} is $\frac{h_i(t)}{h_k(t)} = e^{\sum_{j=1}^{q}\beta_j(z_{ji}-z_{jk})}$. As this ratio is constant in time, individuals i and k are said to have proportional hazards. Using the relationships stated in **Section 1.5**, **(1)** can also be written as

$$S(t) = [S_0(t)]^{e^{\left(\sum_{j=1}^{p}\beta_j x_j\right)}} \quad \textbf{(2)}$$

The PH assumption can be checked for each covariate using a **log-minus-log (LML) plot** of the survivor function i.e. $\log_e[-\log_e S(t)]$ against time t. If the hazards are proportional, the curves generated for the different levels of a covariate, keeping the other covariates constant, should be parallel because $S_1(t)=S_0(t)^c$ from **(2)**, for samples

0 and 1, say (pages 41 & 251, Lee,1992). An LML plot is derived for the HCC patients (***long-acting somatostatin analogue*** treated versus historical controls, **Figure 4.1.5.2.2.**).

Using the Cox approach, the interest is in the proportional factors rather than the baseline hazard. The model is therefore non-parametric with respect to time but parametric in terms of the covariates. The parameter vector **β** is estimated by maximising a partial likelihood. Consider a death at time t_j. The set of all individuals known to be alive just before time t (i.e. have not died or been censored) is known as the **risk set** at time t. The risk of death at t_j for each individual in the risk set is given by **(1)**. Then, conditional on this event (i.e. the death at t_j), the probability that patient i died is $\frac{h_0(t)\exp(\beta x_i)}{\sum_R I(T_R \geq t)h_0(t)\exp(\beta x_R)} = \frac{\exp(\beta x_i)}{\sum_R I(T_R \geq t)\exp(\beta x_R)}$, where $I(T_R>t)$ is an indicator variable and R represents the risk set at time t . Clearly, the expression does not depend on the baseline hazard. The **partial likelihood** for **β** is the product of all terms like the one above over all observed deaths. Following Cox (1975), this likelihood when there are no ties in death times can be written as

$$\prod_i \frac{\exp(\beta^T x_{(i)})}{\sum_{j \in R_i} \exp(\beta^T x_j)} \qquad \textbf{(3)}$$

where $x_{(i)}$ is the value for the individual failing at time $t_{(i)}$, x_j is the value of x for the j^{th} individual and R_i represents the risk set at time t_i. It is the logarithm of this partial likelihood that is maximised. The product is taken over the individuals for whom death times have been recorded. It is assumed that censoring is independent and uninformative (the uninformative censoring meaning that the likelihood for observations censored in [t, t+δt] does not depend on **β**). The technical term 'partial likelihood' refers to the fact that the component terms are derived conditionally on the times that deaths occurred and the composition of the risk set at these times. The times themselves are not used, but the ranked (i.e. ordered) death and censoring times determine the composition of the risk sets. In the sense that it is conditional on the risk sets, the method of partial likelihood is similar to the log rank test described in **Section 2.1.1.4**. If $h_0(t)$ is restricted by a parametric assumption (e.g. if it is constant, as in the exponential model), then the partial likelihood has to be modified accordingly (Cox, 1975).

SPSS fits the Cox model by maximizing the partial likelihood using a Newton-Raphson procedure and appears to use Breslow's approximate likelihood (Breslow, 1974, pg65 Collett, 1994) to cope with tied observations (pages 280 &289, Collett, 1994). When using S-PLUS, Efron's approximation to the likelihood (Efron 1977) will be used to deal with ties for the data analysed in the present study. The Efron and Breslow methods are equivalent when there are no ties in death times, as censoring is assumed to occur after death when there are censored observations at a death time (as mentioned in **Section 2.1.1.1.**). The Efron approximation is much more accurate than the Breslow method when dealing with tied death times (being closer to the appropriate likelihood function) although in practice the two methods often give similar results.

For a given set of data, the larger the value of the **maximized likelihood** (the likelihood function value when parameters are replaced by their maximum likelihood estimates), $\hat{L}$, the better the agreement between model and data. The maximized likelihood can be computed from **(3)** by replacing the β's by their maximum likelihood estimates under the particular model chosen. In practice, it is more convenient to use $-2\log\hat{L}$, minus twice the logarithm of the maximized likelihood, which will always have a positive value. $\hat{L}$ (and therefore $-2\log\hat{L}$) is not useful on its own, as its value depends on the number of observations in the data set. Therefore, $-2\log\hat{L}$ is used in making comparisons between different models fitted to the same data. Two competing models, one with p covariates and the other with p+q covariates (e.g. with sex and age included or not included), may be compared by testing whether the additional q parameter values are significantly different from zero. Under the null hypothesis that they are not, the following hypothesis test can be used: $-2\log_e\left\{\frac{L_p(\hat{\beta})}{L_{p+q}(\hat{\beta})}\right\} \sim \chi^2_q$. This is the **likelihood ratio test**.

For a single unknown parameter β, the **score test statistic** is $\frac{\{u(0)\}^2}{i(0)}$ and the **Wald test statistic** is $\hat{\beta}^2 i(\hat{\beta})$ where u(β) is the efficient score for β, $u(\beta) = \frac{d\log L(\beta)}{d\beta}$ and

i(β) is the observed information function, $i(\beta) = -\left\{ \frac{d^2 \log L(\beta)}{d\beta^2} \right\}$. Each of these statistics has an asymptotic chi-squared distribution on 1 df under the null hypothesis that β=0. The Wald statistic is equivalent to the statistic $\frac{\hat{\beta}}{se\hat{\beta}}$, which has an asymptotic normal distribution. These tests can be generalized to the case of p unknown parameters using a p-component vector β, and a matrix of partial second derivatives (Hessian matrix). Further details of the mathematics for the multiparameter case can be found in Collett (pg 321-2, 1994). For the data sets analysed in the present study, in order to determine which combination of variables provides the most information, forward and backward **stepwise selection procedures** were used in which variables were added and deleted according to the pre-specified cut-off criteria of p-values of 0.05 and 0.1 for entry and removal respectively. Using SPSS, covariates were tested for entry into the model one by one using the significance level of the score statistic. After each entry, the variables already in the model were tested for removal based on the significance of the Wald statistic. To test the overall model (in SPSS and in S-PLUS), the difference between minus twice the log likelihood for the baseline model (in which all β's equal zero) and the present model was computed and the likelihood ratio test described above was used.

SPSS syntax details for Cox models are provided in **Appendix A I.** The Cox PH regression model is widely used in the present study, with Cox modelling being undertaken mainly using SPSS but also using S-PLUS (see **Appendix A II**). For binary X, $\exp(\beta_i)$ is the ratio of the hazard function of a subject with X=1 to that for a subject with X=0 e.g. if X is treatment and β is the logarithm of the ratio of the hazard of death at time t for treated versus untreated patients. If $e^{\beta} > 1$, the conclusion is that a treated person has a greater risk of death at any time than an untreated person. For a two-level characteristic, the **relative risk (RR)** is defined as the ratio of the estimated hazard for a case with the characteristic to that for a case without it (pg 296, Norušis/SPSS Inc, 1994). This definition can be extended to categorical variables with more than two categories, when contrast variables are set up to compare levels of the categorical variable. Details on setting up contrast variables for categorical covariates are given in **Appendix B.** For a continuous variable X_i, the output $e^{\hat{\beta}_i}$ can

be interpreted as the hazard ratio corresponding to a change of one unit in X_i i.e. every unit increase in the i^{th} covariate X_i increases the risk of dying by the multiplicative factor $\exp(\beta_i)$.

2.1.2.2. Checking the Cox PH model

2.1.2.2.1 Residuals

As with other regression models, residual and diagnostic plots can be examined for outliers and lack-of-fit of the survival model. With survival data, however, there are restrictions on the use of residuals and interpretation of these residuals may be much more difficult than for those in linear models. Three common types of residuals used in survival analysis are:

a) Cox-Snell residuals (Cox & Snell, 1968)

b) martingale residuals (Barlow & Prentice, 1988, Therneau et al, 1990)

c) partial residuals (also known as Schoenfeld or score residuals, Schoenfeld, 1982)

The **Cox-Snell residual** for the i^{th} individual is given by $r_{CSi} = \hat{H}_0(t_i)e^{\hat{\beta}' x_i}$, where $\hat{H}_0(t_i)$ is the estimated cumulative baseline hazard function at time t_i, the observed survival time of individual i. Therefore, $r_{CSi} = \hat{H}_i(t_i) = -\log \hat{S}_i(t_i)$, where $\hat{H}_i(t_i)$ and $\hat{S}_i(t_i)$ are the estimated values of the cumulative hazard and survivor functions of the i^{th} individual at time t_i. If T is the random variable associated with the survival time of an individual and S(t) is the corresponding survivor function, then the random variable Y=-log S(t) has an exponential distribution with unit mean, irrespective of the form of S(t) (see Collett, 1994, pg 151 for proof). As the Cox-Snell residuals are estimates of $-\log S(t_i)$, they should have an approximate unit exponential distribution. If the observed survival time of the individual is right-censored then the corresponding value of the residual is also right-censored. The residuals then form a censored sample from the unit exponential distribution and so are expected to have a mean and variance of 1 if the fitted model is correct. One way to assess this is to compute the Kaplan-Meier estimate of the of the 'survivor function' of these values, treating those residuals from censored observations as being censored themselves. If the log-cumulative hazard plot of the residuals $\log\{-\log \hat{S}(r_{CSi})\}$ plotted against log r_{CSi} is approximately a straight line with slope 1 and intercept 0, this indicates that the fitted survival model is appropriate (e.g. **Figure 4.1.1.4, Figure 4.1.1.16**). It should be

noted, however, that for small samples this unit exponential distribution approximation is not at all reliable.

Martingale residuals are formed by taking the difference between the death/event indicator δ_i and the Cox-Snell residuals: $r_{Mi} = \delta_i - r_{CSi}$. They can be thought of as being the difference between the observed number of deaths for individual i in the time interval $(0,t_i)$ and the corresponding estimated expected number on the basis of the fitted model. Martingale residuals can be used to assess whether any particular patients are poorly predicted by the model (with large negative or large positive values indicating a lack of fit). They can also be used together with continuous covariates for assessing the functional form required for the covariate (i.e. does it need transforming?) with a random scatter about zero indicating that the covariate form is adequate (e.g. **Figure 4.1.1.3**). The patterns in plots for categorical variables are, however, often impossible to interpret. Martingale residuals are computed for each subject separately for each variable and they focus on the difference between the covariate values at the failure time of the subject who dies and the covariate means of the corresponding risk set. As these residuals have zero expected mean and are asymptotically uncorrelated, they can be plotted against the time ranks of all individuals (whether they fail or are censored). Any changes in variability or trends are taken as indication of departure from the proportionality assumption. These residuals are difficult to interpret, however, as they are not symmetrically distributed about zero. They take values between $-\infty$ and 1 (and these residuals for censored observations are negative).

Disadvantages of both the Cox-Snell and martingale residuals include the facts that they depend heavily on the observed survival time and they require an estimate of the cumulative hazard function (pg 154, Collett, 1994). These disadvantages are overcome when considering **score residuals**. These residuals do not depend on time so they may be plotted against time to assess the appropriateness of the proportional hazards assumption (e.g. age against time, **Figure 4.1.1.15**). They are plotted only for uncensored cases. If the PH assumption holds, they are expected to be fairly evenly distributed about zero. For each individual there is a set of values of residuals, one for each explanatory variable included in the fitted Cox regression model. The score

residual r_{Sji} is the difference between the j^{th} explanatory variable and a weighted average of the values of the explanatory variable over individuals at risk at the death time of the i^{th} individual. The residuals sum to zero, the expected value of r_{Sji} is zero in large samples and they are uncorrelated with one another. Those individuals who are unlikely to die at time t_i, relative to those who are at risk of death at t_i will have small values of the score residuals and vice versa for those who are more likely to die (pg 155-6, Collett, 1994).

In **SPSS**, both Cox-Snell residuals (these are simply the estimated cumulative hazard function) and partial residuals for each explanatory variable are provided. Also, martingale residuals can be constructed from the estimated cumulative hazard function: if the status indicator is coded 0 for censored cases and 1 for uncensored cases, the martingale residuals are the estimated cumulative hazard function subtracted from the status indicator (page 308, SPSS Inc, 1997). It should be borne in mind, however, that the use of residuals for model checking is more informative in large data sets. The use of residuals in the present study is made for the cirrhosis models for time to decompensation (n=306 patients) and survival (n=307 and n=138 patients presenting with compensated and decompensated cirrhosis respectively).

2.1.2.2.2 Time-dependent covariates

One way of checking whether the PH assumptions are violated (i.e. that the linear component of the model varies with time) is by the addition of a **time-dependent variable** to the model. Altman & de Stavola (1994) state that the term 'time-dependent' may incorrectly imply time dependency of the coefficients rather than the covariates themselves and it may therefore be preferable to use the term ' updated measurements of the covariates'. As the term 'time-dependent' is more widely known, however, it will be used here. There are two types of time-dependent variables: internal variables and external variables. **Internal time-dependent variables** require the survival of the subject to whom they refer in order to exist. An example of an internal time-dependent variable in the hepatoma data set, whose value may be recorded on a regular basis over the period of a clinical trial would be the size of the tumour. This type of variable can be incorporated into survival analysis models, the idea being that more recent values of tumour size may provide a better indication of future life-expectancy than the value at the time origin. **External time-dependent**

variables are either pre-determined, e.g. the age of the subject, or vary independently of the survival process, e.g. air pollution. Two internal time-dependent variables are considered in the present study: the occurrence of decompensation and the occurrence of HCC in cirrhosis patients (**Section 4.1.1.**).

The Cox regression model in which some of the explanatory variables are time-dependent, where the value of the j^{th} variable for the i^{th} individual at time t is written as $x_{ji}(t)$, becomes: $h_i(t) = h_0(t)e^{\sum_{j=1}^{p}\beta_j x_{ji}(t)}$. In this model, the baseline hazard function $h_0(t)$ is interpreted as the hazard function for an individual for whom all the variables are zero at the time origin, and remain so through time. The relative hazard is now time-dependent, as the values of the variables depend on the time t. Therefore, the model is no longer a proportional hazards model.

SPSS has the capability to allow time-dependent variables to be included in the Cox model. The regression coefficients β_j, however, have different interpretations in the time-fixed and time-dependent models. In the time-fixed model, the coefficients β_j represent the effects on the hazard, and therefore on survival, of the entry values of the covariates. In the time-dependent model, however, the β_j represent the effect that the covariates have have at entry and at any time after entry, implying a constant effect over time. Therefore, an assumption underlying the time-dependent model is that the effects of the covariates are time-invariant. If this assumption holds, it is expected that the coefficients in the time-fixed model will be smaller in absolute value than the coefficients in the time-dependent model because of the time decay effect of entry values (Altman & De Stavola, 1994). SPSS syntax details for time-dependent models are provided in **Appendix C.**

It should be noted that if the main aim of the analysis is to assess possible treatment effects in a clinical trial (e.g. treatment with **UDCA, plaquenil, and short- and long-acting somatostatin analogues**), then the time-fixed model is the initial model of choice. It should always be examined and following this, perhaps, a time-dependent model fitted either to check the PH assumption or to extend the model. The reason for this tentative approach to time-dependent modelling is that a treatment effect may be **masked** by relevant time-dependent prognostic variables (if the path of the covariates

is affected directly by treatment) i.e. even though the treatment has an effect on survival, the effect may not be detected by the model (Altman & de Stavola, 1994). An example of the masking effect is seen in a published time-dependent model applied to PBC data. In this model, the time-dependent prognostic variables, and in particular the bilirubin variable, carry the azathioprine treatment effect, resulting in a non-significant therapy effect, whereas the treatment had previously been shown to have a significantly beneficial effect on survival (Christensen et al, 1993). The time-dependent modelling approach is more appropriate if the purpose of a study is to model the evolution of a disease. In the present study, time-dependent modelling is considered in the context of the natural history of cirrhosis (**Section 4.1.1**).

In order to investigate whether the effect of one or more covariates varies with time (which would indicate that a model with time-varying *coefficients* should be used), the data set can be split into smaller sets by censoring follow-up at various time points (Altman & de Stavola, 1994) and fitting the chosen model to the nested data sets. Trends or variations in the estimated coefficients and their significance indicate departures from the PH assumption in the time-fixed model and this method can also be used in the time-dependent model, with drifts in the estimated coefficient of a model indicating possible violation of the assumption of constant effects (Altman & de Stavola, 1994). This approach is employed in the present study for the estimated survival times for decompensated cirrhotics data with follow-up censored at 3 years, 5 years and 7 years (results in **Section 4.1.1**).

2.1.2.2.3. Testing for the presence of informative censoring

The methods used in this thesis for the analysis of censored survival data are only valid if censoring is non-informative i.e. the censoring is not related to any of the factors associated with the actual survival time (informative censoring is defined in **Section 1.5**). One way to examine the possibility of informative censoring is to plot observed survival times against values of explanatory variables, and distinguish the censored observations from the uncensored observations in the plot. If a pattern is present, such as more censored observations for a particular range of the explanatory variable, or at an earlier time on one treatment than the other, then there is a indication of informative censoring. More formally, a linear logistic model could be used to model a binary censoring response (0/1) and estimating whether particular

explanatory variables lead to significant changes in the deviance when included in the model, thus indicating violation of the assumption of non-informative censoring (pg 274, Collett, 1994). If informative censoring is in fact present, there is no satisfactory way to compare groups of patients. The presence of informative censoring was tested for in the cirrhosis data set by plotting sex against survival time, distinguishing between censored and non-censored values (**Figure 4.1.1.17**) and also using logistic regression techniques (**Section 4.1.1.**).

2.1.3 Sample size and the number needed to treat (NNT)

Calculation of the sample size required to detect a relevant effect in a prospective study, as well as calculations to assess the power needed to detect a known effect with the available data in a retrospective study, form an important aspect of any clinical investigation. Sample size and power formulae were developed by Schoenfeld (Schoenfeld, 1981) for survival analyses with randomized treatment comparisons in which it is assumed that there is a factor of interest (e.g. treatment) Z_1. Schoenfeld's formula represents the total number of patients required to detect a relative risk between groups defined by Z_1 i.e. $\theta=\exp(\beta_1)$ under the null hypothesis H_0: $\beta_1=0$ versus H_1: $\beta_1=\ln\theta$ with significance level α and power $1-\beta$. **Shoenfeld' s formula**, for a 2-sided test with significance level α and power $1-\beta$, is: $N=\frac{(u_{1-\frac{\alpha}{2}}+u_{1-\beta})^2}{(\log_e\theta)^2\psi(1-p)p}$,

where u_γ = γ quantile of standard normal distribution, ψ is the probability of being uncensored and $p=\text{Prob}(Z_1=1)$. For example, for a power of 0.8 with $\alpha=0.05$, $u_{1-\alpha/2}=1.96$ and $u_{1-\beta}=0.84$. It is assumed that the probability of censoring $1-\psi$ is nearly identical under H_0 and H_1. Schoenfeld's formula is used here in the context of assessing the ***plaquenil*** treatment effect (see **Section 4.1.2**).

Another aspect that may be considered in treatment comparisons is the **number needed to treat (NNT),** or alternatively **NNTB (number needed to treat to benefit)**, which is defined as the number of patients who need to be treated (with the new treatment) to prevent one additional event (i.e. to achieve one more success than on the old treatment). The NNT concept was originally introduced 10 years ago (Laupacis, 1988, cited in Altman, 2000). In survival analysis, this number can be calculated at any time point after the start of treatment, although there is no **single**

NNT. The NNT is always at least 1.0 and takes its minimum value when the proportions of successes are 0 on the old treatment and 1 on the new treatment (i.e. the new treatment is always effective so all patients survive with it, otherwise all patients die). An estimate of the survival probability in each group at the particular time point ($\mathbf{S_a}$, say, for the active treatment group and $\mathbf{S_c}$ for the control group) and corresponding standard error (s.e.) can be used to calculate the NNT.

The **absolute risk reduction, ARR**, is defined as $\mathbf{S_a}$-$\mathbf{S_c}$ with 95%C.I. ARR±1.96se(ARR), with $se(ARR) = \sqrt{\frac{S_a^2(1-S_a)}{n_a} + \frac{S_c^2(1-S_c)}{n_c}}$ where n_a is the number of patients at risk at the particular time point in the active treatment group and n_c is the corresponding number in the control treatment group. If the standard errors of $\mathbf{S_a}$ and $\mathbf{S_c}$ are known, then se (ARR) can be calculated as $\sqrt{\{[SE(S_a)]^2 + [SE(S_c)]^2\}}$. Then, $NNT = \frac{1}{ARR}$ with 95%C.I. $\left(\frac{1}{A_u}, \frac{1}{A_L}\right)$ where A_L and A_U are the lower and upper confidence limits for ARR. This method of calculating the NNT was used in the present investigations. An alternative way of calculating the NNT involves the use of the estimated hazard ratio and corresponding s.e. (perhaps from a Cox model) and the estimated survival probability for the control group at that time (Altman & Andersen, 1999). Then $NNT = \frac{1}{[S_c(t)]^h - S_c(t)}; h = e^b$ where h is the hazard ratio and b is the regression coefficient. The 95% C.I. for this NNT can be found by replacing h by the two limits of the 95%C.I. for h; if the C.I. is not given explicitly, it can be calculated using $e^{b-1.96se(b)}$ to $e^{b+1.96se(b)}$. This C.I. may be too narrow, however, as it ignores imprecision in the estimate of $S_c(t)$ (Altman & Andersen, 1999).

When there is no treatment effect, the ARR is 0 and the NNT is infinite. This causes difficulty with interpretation of the C.I. for non-significant treatment effects, which will include a negative limit and may seem not to include the best estimate. A negative number needed to treat for the lower confidence limit is also known as the **number needed to treat to harm (NNTH).** By an NNT of – 10, what is meant is that if 10 patients are treated with the new treatment, one fewer has a positive outcome than if they had all received the standard treatment (or no treatment). Altman (1998)

suggests that the C.I. in the case of a non-significant treatment effect be noted as (NNTH X_1 to ∞ to NNTB X_2), where X_1 and X_2 are positive integers, with the graphical scale on a plot ranging from NNTH=1 to NNTB=1 via infinity. In the present study, the NNT was estimated at 12 months for the HCC patients treated with ***octreotide*** and for those receiving ***long-acting somatostatin*** treatment (**Sections 4.1.5.1 and 4.1.5.2.**)

2.1.4 Accuracy assessment

2.1.4.1. The Brier score and measures of residual variation

In the assessment of prognostic classification schemes, it is important to consider two aspects: firstly, the accuracy of the chosen prognostic classification model and secondly, the accuracy of the model in relation to that of other prognostic schemes. In survival analysis, model assessment and comparisons are generally made in terms of Kaplan-Meier estimates of survival probabilities (**Section 2.1.1.1.** above) and using estimated regression coefficients of survival models (**Section 2.1.2.1.** above). Recently, measures of inaccuracy have been developed which may be calculated to assess the usefulness of estimated patient-specific survival probabilities associated with a prognostic survival classification scheme, where the 'patient-specific' probability of being event-free up to time t is defined as S(t/X=x) = P(T>t/X=x), for a given vector of covariates X=x, observed at t=0 (Graf et al, 1999). A measure of inaccuracy can be calculated that relates the estimated patient-specific survival probabilities to the observed outcome, based on a suitable loss function.

Graf et al (1999) describe a partition of the sample space containing 'risk strata' $X_1,\ldots X_g$ where the membership of a particular stratum is described by a one-dimensional covariate defined by $\tilde{X} = j$ if X ε X_j, j=1,…,g. The corresponding estimated probabilities of being event-free up to time t* for patients in risk stratum X_j are denoted $\hat{\pi}(t^* / \tilde{X} = j)$. One example of data with two risk strata would be a randomized allocation to treatment or placebo and comparison using Kaplan-Meier survival curves. Alternatively, for a Cox model, the survival curve can be estimated for each combination of covariates based on the estimated baseline survival function and estimated model coefficients, so each individual in the sample may have a different estimated survival function $\hat{\pi}(t^* / X_j)$.

One measure of inaccuracy that can be used in the presence of random censoring is the **empirical Brier score**, defined as

$$BS^c(t^*) = \frac{1}{n}\sum_{i=1}^{n}\{(0-\hat{\pi}(t^*/\tilde{X}_i))^2 Ind(\tilde{T}_i \le t^*, \delta_i = 1)(\frac{1}{\hat{G}(\tilde{T}_i)}) + (1-\hat{\pi}(t^*/\tilde{X}_i))^2 Ind(\tilde{T}_i > t^*)(\frac{1}{\hat{G}(t^*)})\}$$

where $\tilde{T}_i = \min(T_i, C_i)$ and $\delta_i = Ind(T_i \le C_i)$, i=1,…,n (as in **Section 1.5**) and the censoring time C, is distributed according to G(t)=P(C>t). $\hat{G}(t)$ represents the Kaplan-Meier estimate of the censoring distribution G, based on $(\tilde{T}_i, 1-\delta_i)$, which is equivalent to the reverse Kaplan-Meier process in the absence of ties in the data. The Brier score can be calculated at various points in time in order to assess the overall accuracy of the scheme under consideration. The higher the score, the greater the inaccuracy of the model predictions. The greatest time point is chosen so that censoring is not too heavy e.g. median follow-up time (Graf et al, 1999).

The empirical Brier score can be calculated when $\hat{\pi}(t^*/\tilde{X}_i) = \hat{\pi}(t^*) = \hat{S}(t^*)$ is used as a prediction for all patients, $\hat{S}(t^*)$ being the Kaplan-Meier estimate at t*, and is denoted $BS^c_0(t^*)$. Details for hand calculation of the empirical Brier score, as used in the present study, are provided in **Appendix D.**

The measure of **explained residual variation R^2** is defined as $R^2 = 1 - \frac{BS^c(t^*)}{BS^c_0(t^*)}$. If the accuracy of the Cox model is being assessed, then values of R^2 close to 0 indicate that there is no advantage over simple Kaplan-Meier estimates.

In the survival analysis model, the time-to-event itself cannot be accurately predicted. The best that can be done when there is prognostic information available at t=0 is to estimate the probability that the event will not occur until a certain time t*, given the observed covariate information (Graf et al, 1999). Therefore, the measures of inaccuracy are comparisons of the estimated event-free probability of the observed individual outcome. In the present thesis, empirical Brier scores and estimates of the explained residual variation were obtained for Cox PH model predictions of the time to decompensation for compensated cirrhotics (**Section 4.1.1, Figure 4.1.1.7**).

2.1.4.2. The bootstrap and the jackknife

Bootstrap and jackknife techniques are used in the present study to investigate the stability of Cox regression models in terms of the choice of variables included in the model. The aim is to confirm the large-sample approximations for Cox regression in relation to estimates of regression coefficients. The non-parametric bootstrap (Efron 1982, Efron & Tibshirani 1986) and the parametric jackknife (M.H. Quenouille, 1956) are computer-intensive resampling methods. The basic idea behind the bootstrap is that if independent identically distributed observations $X_1,X_2,\ldots,X_n$ are available then available characteristics of the distribution of the X's can be assessed by studying the variability of the estimate across a large number B of bootstrap samples (Altman & Andersen, 1989). The bootstrap samples are obtained by taking samples of size n from the original data using random sampling with replacement. The mean of estimated statistics from the bootstrapped samples approximates the mean of the population and the standard deviation of the estimate approximates the standard error of the statistic as if there had been repeated sampling from the population without replacement (pg823, Sokal & Rohlf, 1995). The statistics considered here are the estimated regression coefficients from the Cox model. The estimated bias is calculated as the difference between the mean of the replicates and the observed values from the original data. Two types of percentile estimates will be used to obtain confidence limits: **empirical percentiles**, which are simply the percentiles of the empirical distribution of the replicates and bias-corrected and adjusted (**BCa**) **percentiles**, which require more computational time than the empirical percentiles but are believed to be more accurate. In the analyses undertaken in the present study, B is taken to be 1000, the recommended minimum number for the empirical percentile limits calculated to be sufficiently accurate (pg 823, MathSoft Inc, 1997), unless otherwise stated. The S-PLUS bootstrap function and an example of its application is provided in **Appendix E**.

In jackknife resampling, a statistic is calculated for n possible samples usually of size n-1, each with one observation omitted (as used here). Jackknife estimates of bias, mean and standard error are calculated in a different way from the equivalent bootstrap statistics. 'Jackknife after bootstrap' is used in S-PLUS to obtain estimates of the variation in the functionals (SE, mean and bias) of the bootstrap distribution and to examine the influence of particular observations on the functionals (see

Appendix E). In the present study, the focus is on the variability of the bias. 'Jackknife after bootstrap' provides standard error (SE) estimates for the bias i.e. the mean of the distribution of biases. Therefore, the SE of the bias is the SE of the mean, and the 'influence' indicates the influence of each observation on the mean. Influence plots (the influence having been calculated using normalized versions of the SE estimates) give an indication of which observations are particularly influential, the criterion being an absolute relative influence greater than 2 (pg 840, MathSoft Inc, 1997). The models investigated using resampling techniques in the present study were the overall Cox survival model coefficients for cirrhotics (**Section 4.1.1.**), the coefficients in the time to decompensation model for compensated cirrhotics (**Section 4.1.1.**) and Cox model coefficients for the PBC patients (**Section 4.1.3.**).

2.1.4.3. Cross-validation techniques

The **cross-validation**, or **leaving-one-out**, **technique** is applied in the present study using SPSS to internally validate the discriminant analysis model derived for classification of the peritoneal effusions of patients with ascites (the patients described in **Section 3.4** below). Cross-validation is used to obtain an estimate of the misclassification rate when the model is applied to new data. Using this technique, each of the cases is left out in turn, the discriminant functions are calculated based on the remaining cases and the omitted case is subsequently classified. As the omitted case has not been used in the calculations, the misclassification rate obtained is thus presumed to be less biased than the misclassification rate obtained directly from the discriminant analysis classification model.

2.1.5. Comparisons with the Mayo prognostic model

2.1.5.1. The Mayo model

In the medical literature, one frequently finds articles in which observed survival times of groups of patients are compared with survival times predicted by previously developed and validated statistical models. One such extensively referenced model is the Cox proportional hazards "Mayo model" developed at the Mayo Clinic, U.S.A. (from a database on 312 patients referred to the clinic between January 1974 and May 1984) in order to improve the selection of patients for and timing of liver transplantation (Dickson et al, 1989). The end-point used in the original model was death from any cause (with transplant patients being censored at the date of

transplantation). The model was cross-validated on 106 Mayo Clinic patients, who were subsequently incorporated into the analysis, the final model parameters being derived from 418 patients, 25 (6%) of whom underwent liver transplantation. Extra-mural validation was also undertaken using 176 PBC patients from Boston and Texas, U.S.A.(Grambsch et al, 1989). Five variables are combined in the model to obtain a risk score (R) for each patient, which has the following form:

R=0.871$\log_e$(bilirubin in milligrams per decilitre)-2.53$\log_e$(albumin***** in grams per decilitre)+0.039(age in years)+2.38$\log_e$(prothrombin time in seconds)+0.859(oedema score).

*****Albumin was measured by serum protein electrophoresis.

Larger values of **R** indicate a higher risk (i.e. poorer prognosis). If S(t,X) denotes the probability that a patient with risk factors given by X=(X_1,…,X_p) and with risk score R will still be alive t years later and it is assumed that there is a known survival function $S_0(t)$ for individuals with risk score R_0 then, from the PH assumption, $S(t,X)=\{S_0(t)\}^{\exp(R-R_0)}$ (from equation **(2)**, **Section 2.1.2.1.**). The underlying survival function for the Mayo model with R_0=5.07 is provided in **Table 2.1.5.1.1.** below.

Table 2.1.5.1.1. Underlying survival function for the original Mayo model

t (years)	1	2	3	4	5	6	7
$S_0(t)$	0.970	0.941	0.883	0.833	0.774	0.721	0.651

$S_0(t)$ gives the survival probabilities for a patient with risk score 5.07, the mean of the combined Mayo data set.

*Adapted from **Table 4**, pg 6, Dickson et al, 1989*

An updated version of the Mayo model has also been applied in recent PBC studies in which the end-point is considered to be either death or liver transplantation (Lindor et al, 1996, Poupon et al, 1999) as compared to the original model in which patients undergoing liver transplantation were censored at the date of transplantation. Treating transplantation patients as being censored on the date of transplantation has been disputed as violating the assumption of random censoring. The incorrect censoring occurs because patients who undergo transplantation are known to be at a higher risk of death than other patients (as discussed by Bonnand & Poupon, 1996) so censoring time cannot be considered independent of end-point. Independence is one of the underlying assumptions in the proportional hazards model-fitting procedure. The

updated Mayo model, although differing only slightly from the original model because only 6% of patients underwent transplants, takes account of transplantations by treating both liver transplantation and deaths as events, keeping the model coefficients fixed (Lindor & Therneau, 1996). Any bias in using the updated model is therefore "in favour of" the Mayo model, in that those patients who undergo liver transplantation are not on the verge of dying (although treated as such in this model). In the present investigation, the updated Mayo model was used to create a 'simulated control group'. The model remains the same as the original model apart from the underlying survival function, which is displayed in **Table 2.1.5.1.2.** below.

Table 2.1.5.1.2. Underlying survival function for the updated Mayo model

t (years)	**1**	**2**	**3**	**4**	**5**	**6**	**7**
$S_0(t)$	0.970	0.938	0.866	0.805	0.737	0.682	0.588

$S_0(t)$ gives the survival probabilities for a patient with risk score 5.07, the mean of the combined Mayo data set.

Adapted from pg 1783, Lindor & Therneau, 1996

Since its development in 1989, the Mayo model has been widely used in comparing actual with predicted survival times in groups of patients undergoing liver transplantations (Markus et al, 1989), in UDCA-treated patients (Lindor et al, 1996, Poupon et al, 1999), in non-U.S. PBC patients (Krzeski et al, 1999) and also in sufferers of PBC in the community (Kim et al, 2000). The main advantage of the Mayo model over other similarly developed models (European model, Yale model, Oslo model, Glasgow model, Australia model, cited in Wiesner, 1998) is that it does not necessitate liver biopsy (the justification by the authors not including biopsy results as a variable in the prognostic model being the strong correlation between stage and Mayo risk score) and hence can be considered a 'bed-side' application. The measurements used in the Mayo model are inexpensive, non-invasive, and universally available. The European model was constructed using similar entry criterion to the Mayo model whereas the Yale model used the estimated date of onset of PBC as initial time point although the variables in the model were measured at the date of diagnosis. Survival comparisons were made between the updated Mayo model predictions and the Cretan UDCA-treated PBC data (**Section 4.1.3.**).

2.1.5.2 Graphical and statistical comparisons

Using the available Cretan single-time point PBC data ('single time point' meaning that updated measurements were not available), the baseline characteristics of 114 patients undergoing UDCA treatment are compared with those of Mayo model predictions (Table 3, pg 5 Dicksen et al, 1989). Actual survival was estimated using the Kaplan-Meier PL estimator (as described in **Section 2.1.1.1.**). Predicted survival was calculated for each patient using the Mayo model. Subsequently, graphical comparisons were made of the Kaplan-Meier PL estimate of the survival curve with that predicted by the Mayo model, the latter having been obtained using the **direct-adjusted method** of Thompsen et al (1991). The mean Mayo model curve is the average of the per-patient survival curves that are predicted by the Mayo model. In mathematical terms, the estimated survival probability at time t for patient i with covariates z_i can be written as $\hat{S}_i(t)=\hat{S}(t,z_i)$ and an estimate of the mean survival function $S_m(t)=\Sigma S_i(t)/n$ of the n PBC patients with covariate vectors $z_1,\ldots,z_n$ (analogous to that given by the Kaplan-Meier curve) is $\bar{S}(t)=\frac{1}{n}\sum_{i=1}^{n}\hat{S}(t,z_i)$. In applying the direct-adjusted method, i.e. in assuming that the Kaplan-Meier survival curve may be regarded as an estimate of the mean survival curve obtained as the average of the individual-specific survival curves, it is assumed that the potential follow-up times are the same for all patients (M Voeth, 1992). The averaged Mayo-model predictions were available only at yearly intervals (and only for 7 years following diagnosis) so linear interpolation was used.

In order to test for differences between the survival predicted by the Mayo model for untreated patients and the actual survival of our UDCA-treated patients, the **one-sample log-rank test** was applied (Woolson, 1981). The one-sample log-rank test is often applied to indicate possible differences in survival times between observed survival times and Mayo clinic predictions using the Mayo-model predicted survival curve for each patient as a control for that patient (Markus et al, 1989). Although the one-sample log rank test is the method most widely used in recent PBC medical literature to compare with Mayo model predictions (Poupon et al, 1999 E1, Krzeski et al, 1999, Markus et al, 1989 and others), the use of such tests may not be completely satisfactory in this context as the mean survival function is random, not fixed as assumed by the test.

2.1.6. Incidence rates

The term **incidence** refers to new cases of disease occurring among previously unaffected individuals. The rate of occurrence of an event in a population is the number of events that occur during a specified time interval divided by the total amount of observation time accumulated during that interval. When estimating an **incidence rate**, the 'events' are new cases of disease occurring among disease-free individuals. The denominator is the sum of the length of time during the specified interval that each member of the population was in the study and disease-free (pg42-43, Breslow & Day, Vol I, 1980). The incidence rate is normally expressed as the number of person-years of observation. Crude HCC incidence rates in Cretan cirrhotic patients were estimated in the present study (**Section 4.1.1**).

When the survival analysis techniques described above (**Sections 2.1.1 and 2.1.2**) are used in modelling the time to incidence, the hazard rate **h(t)** is in fact the instantaneous incidence rate (i.e. the incidence rate defined for each instant t of time) and the **cumulative incidence rate** is the sum of the hazard rates over the time interval. Estimates of hazard, and hence, cumulative HCC incidence rates were obtained by Cox PH regression analysis for cirrhotic patients. The cumulative probability of remaining HCC-free at specific time points was estimated using the Kaplan-Meier method (**Table 4.1.1.14**).

2.2 Multivariate techniques to distinguish between ascites diagnostic groups

2.2.1. Recursive partitioning methods

Tree-based modelling is an exploratory technique used to uncover structure in data. In the present context, classification trees were constructed as an alternative to logistic regression modelling. The basic idea in tree-based modelling is to derive a set of decision (or classification) rules using a procedure known as **recursive partitioning**. This technique involves the formation of subgroups, within which there is homogeneity and between which the outcomes being distinct. The procedure may be seen as a kind of variable selection that handles interactions between variables automatically (pg 329, Venables & Ripley, 1994). The path found in the decision tree in graphical form is followed from the top node (called the 'root') to a terminal node (called a '**leaf**'), according to the rules, which are known as '**splits**', found at the

interior nodes. The first split is the most important predictor. In the present study, the classification tree procedure was used to distinguish between 27 patients with malignant ascites and 23 patients with ascites caused by cirrhosis, using ascitic fluid to serum ratios of various biochemical parameters (details of the particular biochemical measurements considered are provided in **Section 3.4**). S-PLUS version 4.5 was used for the recursive partitioning techniques employed. The final tree obtained is depicted in **Figure 4.2.1.1**.

In S, the tree can be seen as providing a probability model. At each node **i** there is a probability distribution $\mathbf{p_{ik}}$ over the classes (here there are two classes so k=1,2). Each case in the data set is assigned to a leaf so at each leaf **i** there is a random sample $\mathbf{n_{ik}}$ from the multinomial (or binomial) distribution specified by $\mathbf{p_{ik}}$. It is the **deviance** (likelihood ratio statistic) which is used to determine which partition of a node is 'most likely' given the data (pg 413, Chambers & Hastie, 1993), where the deviance of the tree is $\mathbf{D=\Sigma D_i}$ the summation being over nodes **i**, $\mathbf{D_i=-2\Sigma n_{ik}log(p_{ik})}$, summed over the classes **k**. The estimated proportions are given by the observed numbers of each class divided by the total number at the node. The reduction in deviance when a node is split in two gives a measure of the value of a split. The partitioning process takes the maximum reduction in deviance over all allowed splits of all leaves to choose the next split. The tree construction process thus uses a 'one step lookahead' (pg 332, Venables & Ripley, 1994) i.e. the next split is chosen each time in a optimal way. Further theoretical details are provided in Venables & Ripley (1994) and Chambers & Hastie (1993). For the ascites patient data, the minimum split size and minimum node size (i.e. minimum number of patients at the node) were initially specified to be the S-PLUS defaults of ten and five respectively but were subsequently relaxed to five and two respectively due to the small numbers involved. Such small split and node sizes are bound, however, to have an effect on the overall accuracy of the model.

2.2.2. Discriminant analysis

The basic problem solved by discriminant analysis in the present study is that of separation of patients with ascites into three diagnostic groups according to the nature of their peritoneal effusions, using patient samples from the three populations

representing disease state (transudate, malignant exudate, non-malignant exudate). A description of the ascitic patient groups is given in **Section 3.4**. Separation is based on a combination of biochemical parameters, given that classification of a patient with unknown disease status into one of the three groups is not possible based on only one biochemical measurement. With one biochemical measurement, the means of the three distributions may not be identical but the distributions may overlap considerably. The aim was to consider the combined effect of all biochemical variables to discover which combination of variables leads to the maximum discrimination between the three groups. Linear discriminant analysis was applied to find a rule to discriminate between the three distinct diagnostic groups, with $\mathbf{n_j}$ (n_1=23, n_2=13, n_3=25) individuals in the $\mathbf{j}^{th}$ group (j=1,2,3), each individual having been measured on **v** variables, $\mathbf{x_1,\ldots,x_v}$. It was assumed that the original classification into groups is made independently of the x variables and is known a priori.

The simplest case in discriminant analysis occurs when there are only two populations. With two populations, the basic strategy is to form a linear combination of the variables $\mathbf{z= b_1x_1+\ldots+ b_vx_v}$, known as **Fisher's linear discriminant function**, and then to assign a new individual either to group A or group B on the basis of the value of z obtained for that individual. Values of $\mathbf{b_1,\ldots,b_v}$ are chosen to provide maximum discrimination between the groups, the idea being to make the variation in z between the groups much greater than the variation within the groups. Therefore, the ratio Δ^2 is maximized, where Δ^2=(mean z_A- mean z_B)^2/(variance of z within groups). A completely symmetrical rule (assuming equal prior probabilities) would be to use the mean z_0 of mean z_A and mean z_B as the allocation cut-off and so, if mean(z_A)> mean(z_B), allocate an individual to A if $z>z_0$, otherwise to B. When there are more than two populations, the above procedure is generalized (in SPSS) to maximization of the ratio of the sum of squares (SS) between groups to the SS within groups. This leads to calculation of the eigenvalues (also called latent roots) of a matrix. The solution which corresponds to the largest eigenvalue gives the linear function coefficients which maximise the ratio of the SS, and is called the **first canonical variate** or **first canonical discriminant function** (the latter term being given in SPSS output). The second canonical variate gives the next highest ratio, subject to the condition that it is uncorrelated with the first, etc. The number of canonical variates is min(v, g-1), where g is the number of groups. Fisher's linear

discriminant function for two groups may therefore be viewed as the first and only canonical variate.

With our three populations, discrimination takes place in the two-dimensional space defined by the two canonical variates and an individual is allocated to the group for which the distance between the individual's data point and the group mean in the x-y plane is least (pg 342, Armitage & Berry,1987). The average score for a group is called the **group centroid** in SPSS. Further algebraic details for linear discriminant analysis can be found in Everitt & Dunn (pg 238, 1991). There are two assumptions in applying linear discriminant analysis. The first is that the covariance matrices for all groups are equal (otherwise a quadratic discrimination function may be more appropriate). The second assumption is that the variables are from a multivariate normal (MVN) distribution, although if violation is not too severe discriminant analysis may still be applied. Checking the distributions of individual variables may provide a clue as to whether an MVN distribution is likely, as if there is an MVN distribution, the individual variables will be normally distributed (although the opposite is not necessarily true). The equality of covariance matrices assumption was tested in SPSS using **Box's M test**, although the test is sensitive to departures from MVN (i.e. tends to call covariance matrices unequal if the normality assumption is violated). Only subjects with complete data were included in the discriminant analysis. A backwards stepwise selection procedure was used in SPSS, with the minimization of **Wilk's lambda**, to determine the most influential variables in the discrimination process. Equal priors were assumed i.e. it was assumed that the probability of a patient belonging to any of the three groups is equal. A territorial map was constructed to display group separation on the basis of the first two linear discriminant functions (**Figure 4.2.2.1**).

2.3 Model validation

The idea of validating a prognostic model is usually taken to mean establishing that it works for patients other than those from whose data the model was derived (Altman & Royston, 2000). The performance (prediction accuracy) of a prognostic model may be assessed using a variety of approaches, as have been described in the previous paragraphs, including comparisons of observed and predicted event rates. An issue of semantics occurs in the use of the term 'valid' as the word 'validity' is a psychometric

measurement method term meaning "Does a measurement method measure what it is supposed to?" With psychometric data, a correlational approach is usually taken to assess within-subject to between-subject variation. In model fitting, however, such an approach is not appropriate as the main issue is the quality of predictions for individuals or groups of subjects. The term 'validated' means something wider than mere performance evaluation. A statistically validated model may be clinically invalid if, for example, there is not enough available intrinsic prognostic information.

Definitions for two types of validated model have been proposed by Altman & Royston (2000):

1) a **statistically validated model**, motivated by the question "With the available factors, is the model the best that can be found?" This model is one that passes the statistical tests, including goodness-of-fit on the original data and unbiased prediction on new data.
2) a **clinically validated model** motivated by the question "Does the model predict accurately enough for the required clinical aims? This model is one that performs well on a new data set, according to context-specific statistical criteria laid down for it.

From the above definitions, a clinically validated model may be statistically invalid (e.g. if there is strong prognostic information, even a biased model may provide a clinically useful separation of patients into prognostic groups) and vice versa (e.g. if the intrinsic prognostic information is too weak the predictions, even if unbiased, will not enable a clinically useful separation). Altman & Royston state that "a clinically validated model is likely to be more useful than a statistically validated one" (even though the first author is a renowned statistician and the article was published in a statistical journal!). The ability to develop a successful model depends on the following features (Altman & Royston, 2000):

a) the potential for accurate prognosis; presumably this is unknown
b) the intrinsic prognostic information in the variables available; this will depend on the physiology of the disease, among other factors
c) the measurement process; some measurements may be more reliable than others
d) the accuracy with which the measurements are converted to predictions

Three main types of validation strategies exist:

1) internal validation: procedures are restricted to a single data set

2) temporal validation: evaluation on a second data set from the same centre
3) external validation: evaluation on data from a different centre

In the present thesis, internal validation of the discriminant analysis models was undertaken using cross-validation techniques (described in **Section 2.1.4.3**) whereas bootstrapping techniques were applied for the survival analysis models (**Section 2.1.4.2).** An estimation method of predictive accuracy that was undertaken for the survival analysis models, where applicable, was the **Brier score,** as described above (**Section 2.1.4.1**) Temporal and external validation were not possible as no other data sets were available in the time span of the present thesis.

3. SUBJECTS

3.1 Cirrhosis patients

The first and largest patient group considered were the 470 cirrhosis patients whose prognostic data were entered into the Gastroenterology Clinic database between the opening of the Clinic and December 2000.

The available prognostic factors at presentation were the following:

- sex (62% male)
- cirrhosis aetiology (see **Section 1.1** for definitions)
    - chronic hepatitis C (HCV, 40%)
    - alcohol (27%)
    - chronic hepatitis B (HBV, 14%)
    - alcohol in combination with viral infection (6%, 31% of these in combination with HBV, 69% in combination with HCV)
    - cryptogenic aetiology (6%)
    - other aetiology (8%)
- age (mean 63 years, s.d. 12.0, ranging from 18 years to 88 years)
- type of decompensation (at presentation or during follow-up)
    - ascites (60%)
    - variceal bleeding (16%)
    - hepatic encephalopathy (5%)
    - other/unknown (17%)
- the occurrence of HCC over the follow-up period (12%)

Prior to the statistical analysis, the records were divided into the following sub-groups according to the circumstances of admission to the Clinic (whether the patient had compensated or decompensated cirrhosis & whether diagnosis was made at presentation or prior to presentation):

1) patients diagnosed when they presented at the Clinic with compensated cirrhosis; 312 subjects
2) patients diagnosed when they presented at the Clinic with decompensated cirrhosis; 98 subjects
3) patients who presented at the Clinic some time after initial diagnosis, presenting with decompensated cirrhosis; 40 subjects (i.e. the initial diagnosis occurred before the Clinic opened)

4) patients who presented at the Clinic some time after initial diagnosis, presenting with compensated cirrhosis; 8 subjects
5) patients who presented at the Clinic some time after both diagnosis and the occurrence of decompensation; 6 subjects
6) patients with unknown diagnosis date and unknown date of presentation to the Clinic; 6 subjects

The demographic characteristics of the Cretan cirrhotic patients at presentation to the clinic are presented in **Table 3.1.1** below. Sixty-three percent of the patients decompensated, either presenting at the clinic with decompensation or decompensating at some later date. Death was classed as being due to **liver failure** if it was associated with the progressive impairment of liver function. The survival times of patients who died from causes independent of the cirrhosis are regarded as right-censored. For those patients whose survival status at the end of the study was unknown, the time from diagnosis to the time at which they were last known to be alive (e.g. at onset of complications) is regarded as a censored survival time.

Using the data from group A in **Table 3.1.1.**, the decompensation-free time and overall survival time of compensated cirrhotics from time of their diagnosis were estimated. There were 306 patients included in the time-to decompensation analysis (of the 312 patients in the data base, 5 patients had missing diagnosis dates and 1 had unknown date of decompensation; these 6 patients were omitted from statistical analysis). There were 150 events (occurrence of decompensation) and 156 censored cases (51%). The median age at diagnosis was 64 years. Follow-up from diagnosis to decompensation ranged from 1 to 136 months. The median follow-up time was 55 months (estimated using reverse censoring).

Table 3.1.1. Characteristics of all 470 cirrhosis patients entered into the Gastroenterology Clinic database, 1989-2000.

Group	Diagnosis of compensated cirrhosis at clinic **A**	Diagnosis of decompensated cirrhosis at clinic **B**	Presents at clinic after initial diagnosis with compensated cirrhosis **C**	Presents at clinic after initial diagnosis with decompensation **D**	Presents at clinic after initial diagnosis and after decompensation **E**	Unknown diagnosis date and unknown presentation date **F**	All patients in database
n	312	98	8	40	6	6	470
Sex							
Male	169	79	5	28	5	5	291
Female	143	19	3	12	1	1	179
Type of cirrhosis							
Alcohol	56	48	2	16	2	4	128
Alcohol+virus	17	7	0	2	0	0	26
Hepatitis B	45	13	0	5	1	2	66
Hepatitis C	151	16	5	12	2	0	186
Κρυψ	13	11	0	3	1	0	28
Other/No aetiology given*	30*	3	1	2	0	0	36
Age at presentation	62.3 (12.0) 12 missing	62.9 (13.0) 1 missing	68.6 (2.8) 1 missing	67.3 (10.7)	61.3 (12.5)	57.8 (3.1) 1missing	62.9 (12.0)
Age at diagnosis	62.3 (12.0)		68.6 (2.8)			57.8 (3.1)	
Decompensate ?							
Yes	154	98	0	40	6	0	298
Type of decompensation							
Variceal bleeding	23	15	0	7	2	0	47
Hepatic encephalopathy	11	1	0	4	0	0	16
Ascites	88	67	0	22	3	0	180
Ascites & hepatic encephalopathy	2	4	0	1	0	0	7
Ascites & ABP	7	4	0	1	0	0	12
Ascites & variceal bleeding	2	1	0	0	0	0	3
Variceal bleeding & hepatic encephalopathy	1	0	0	0	0	0	1
Other	0	1	0	0	0	0	1
Unknown	20	5	0	5	1	0	31
Develop HCC?							
Yes	36	9	0	8	1	1	55

*Includes dual B & C infection in 3 patients

Of the 307 patients diagnosed as having compensated cirrhosis who were included in the overall survival analysis, 70 died and 237 were censored (77.2%). Follow-up from diagnosis to death ranged from 1 to 138 months, with median 56 months. There were

3 losses-to-follow-up, all three occurring at one month after diagnosis (at which point they decompensated). The single end-point of the study was death from liver disease. The 16 patients who died without signs of decompensation were censored at the times of their death i.e. deaths occurring without prior decompensation were assumed to be deaths unrelated to the underlying liver disease. Prognostic survival and decompensation models are presented in **Section 4.1.1.**

Of the 470 patients in the cohort, 144 (31%) presented at the Clinic with decompensated cirrhosis, either at the time of diagnosis (98 patients, group B) or at some date after the initial diagnosis (40 patients, group D). Six of the database entries were of patients presenting after decompensation had already appeared. These six cases were excluded from further analysis. In the decompensated cirrhosis group there were 66 events (i.e. deaths) in total and 72 censored cases (52%). The median age at presentation was 64 years. Follow-up ranged from 1 to 136 months, with median 59 months. There were 27 losses to follow-up after presentation with decompensation (in fact, the 27 patients were lost-to-follow-up within one month of presentation and a further 6 patients actually died within a month of presenting with decompensation). The characteristics of these 27 patients are presented in **Table 3.1.2** below, in which it can be seen that their baseline characteristics in terms of sex, cirrhosis aetiology, age and type of decompensation are very similar to those of the 111 patients who were not lost to follow-up.

In **Table 3.1.3** below are presented summaries by cirrhosis aetiology (alcohol, HBV, HCV, a combination of virus plus alcohol, cryptogenic/other) for compensated and decompensated cirrhotic patients. It can be seen that the baseline characteristics of compensated and decompensated cirrhotics differ between the five aetiological groups. In the compensated cirrhotic group, although the male to female ratio is roughly even overall and for those with cryptogenic cirrhosis, the cirrhotics with alcohol or a combination of alcohol and viral markers as aetiology are overwhelmingly male (96% male alcoholic cirrhotics). There are more males than expected with HBV as underlying aetiology of their compensated cirrhosis (31 observed, 24 expected) and many fewer than expected with HCV (47 observed, 82 expected) under the assumption of independence ($p<0.00001$).

Table 3.1.2. Characteristics of the 27 patients presenting with decompensation who were lost to follow-up within one month of diagnosis and the 111 patients remaining in the study.

		Subjects	
		Lost to follow-up (n=27)	**Remaining in study (n=111)**
Sex	Male	24 (89%)	83 (75%)
	Female	3 (11%)	28 (25%)
Cirrhosis aetiology	Alcohol	15 (56%)	49 (44%)
	Alcohol+virus	1 (4%)	8 (7%)
	HBV	4 (15%)	14 (13%)
	HCV	5 (19%)	23 (21%)
	Cryptogenic	2 (7%)	12 (11%)
	Other/unknown	0	0
Age (mean, s.d.)		61.9 (12.2)	64.7 (12.6)
Decompensation type	Variceal bleeding	3 (11%)	19 (17%)
	Hepatic encephalopathy	1 (4%)	4 (4%)
	Ascites	20 (74%)	69 (62%)
	Ascites+hepatic encephalopathy	1 (4%)	4 (4%)
	Other	1 (4%)	6 (5%)
	Unknown	1 (4%)	9 (8%)

There is also strong evidence of a difference in ages between the groups, with HCV cirrhotics appearing older on average than the other groups ($p<0.0005$). HCV cirrhotics have mean age 66 years, se 0.7 as compared to those with both alcohol and virus as aetiology who have mean age 52 years, se 3.0. In the decompensated cirrhosis group, alcohol is much more common an aetiology than in the compensated cirrhotics (46% versus 18%). The proportion of females in the former group is much lower than in the latter (22% versus 46%). However, very similar patterns in age and sex distributions by aetiology are seen in those presenting with decompensated cirrhosis (**Table 3.1.3.)**. The higher mean age of the HCV cirrhotic patients compared to the HBV patients does not necessarily imply a longer HCV infection or a longer time to development of cirrhosis. In fact, HBV positive patients are known to generally acquire the infection at an earlier age than HCV positive patients (Chiaramonte et al, 1999). The two hepatitis patient groups may therefore have similar average durations

of infection and progression times to cirrhosis, even though the HCV-infected subjects are older on average in both compensated and decompensated cirrhotics.

Table 3.1.3. Patient characteristics by cirrhosis aetiology for 312 patients diagnosed with compensated cirrhosis and 138 patients presenting with decompensated cirrhosis.

		Type of cirrhosis					Total
		alcohol	HBV	HCV	HBV + alcohol/ HCV+ alcohol	cryptogenic/ other	
Compensated cirrhotics							
	N (%)	**56 (18%)**	**45 (14%)**	**151 (48%)**	**17 (5%)**	**43 (14%)**	**312 (100%)**
Sex[1]	Male	54 (96%)	31 (69%)	47 (31%)	16 (94%)	21 (49%)	169 (54%)
	Female	2 (4%)	14 (31%)	104 (69%)	1 (6%)	22 (51%)	143 (46%)
Mean age in years[2] (se)		59 (1.5)	63 (2.1)	66 (0.7)	52 (3.0)	62 (0.7)	60 (2.5)
Decompensated cirrhotics							
	N	**64 (46%)**	**18 (13%)**	**28 (20%)**	**9 (7%)**	**19 (14%)**	**138 (100%)**
Sex[3]	Male	62 (97%)	14 (78%)	12 (43%)	9 (100%)	10 (53%)	107 (78%)
	Female	2 (3%)	4 (22%)	16 (57%)	0 (0%)	9 (47%)	31 (22%)
Mean age in years[4] (se)		60 (1.4)	66 (2.5)	71 (1.8)	59 (5.4)	68 (3.5)	64 (1.1)
Decompensation type[5]							
	variceal bleeding	9 (14%)	2 (11%)	3 (11%)	2 (22%)	6 (32%)	22 (16%)
	hepatic encephalopathy	7 (11%)	2 (11%)	0 (0%)	1 (11%)	0 (0%)	10 (7%)
	Ascites	40 (63%)	14 (78%)	23 (82%)	4 (44%)	8 (42%)	89 (64%)
	Other	8 (13%)	0 (0%)	2 (7%)	2 (22%)	5 (26%)	17 (12%)

[1] chi-squared test statistic(4 df) =87.93, $p<0.00001$

[2] Kruskal-Wallis test statistic (4 df)=26.31, $p<0.0005$

[3] chi-squared test statistic (4 df) =42.449, $p<0.00001$

[4] Kruskal-Wallis test statistic (4 df)=17.96, $p=0.001$

[5] Numbers in each cell were not sufficiently large for a chi-squared test

In addition to estimation of survival times for compensated and decompensated cirrhotics, estimation of the HCC incidence rate in those patients diagnosed at the clinic (with either compensated or decompensated cirrhosis) was undertaken. Here, the end-point was taken to be the occurrence of HCC with deaths prior to HCC being treated as censored values (81cases). Of the 410 patients in this group (groups A and B combined), 45 were diagnosed as having HCC during the study period. In 6 cases, however, the diagnosis of HCC occurred within 1 month of the diagnosis of cirrhosis (4 cases of decompensated cirrhosis and 2 cases of compensated cirrhosis) and in one

case, the date of diagnosis of HCC was unknown. In a further 5 cases, the cirrhosis diagnosis date was unknown (all cases of compensated cirrhosis). These 12 patients were excluded from the analyses. In the present analysis, therefore, 38 (9.6%) of the 398 patients were considered to have developed HCC after the diagnosis at the clinic of cirrhosis. Follow-up ranged from 1 to 138 months with a median follow-up of 51 months. Of the 398 patients, 305 were diagnosed as having compensated cirrhosis whereas 93 had decompensated cirrhosis at diagnosis.

3.2 Patients with PBC

The data consist of prognostic measurements at diagnosis on 114 individuals who were consecutively diagnosed at the Gastroenterology Clinic as having PBC between September 1989 and March 2000. These patients fulfilled the clinical, biochemical, serologic and histologic criteria for PBC. Follow-up was until March 2001. Of the 114 PBC patients in the cohort, 9 had missing information with regard to diagnosis date and one person had missing biochemical measurements. None of these subjects died within the study period. These 10 subjects were completely omitted from the analyses. One patient was lost to follow-up after her initial presentation to clinic; she is considered censored after 1 month for the purposes of the present analysis. The median follow-up time (estimated using reverse censoring) was 64 months, and the follow-up time ranged from 1 to 141 months. There were no liver transplants undertaken at the University Hospital during this time period. The single end-point was death related to the disease. There were 17 deaths during the study period (16% of the 104 patients included in the analysis). At the time of the analysis, 86 of the 104 patients were still alive. All patients were administered UDCA treatment following diagnosis. Dosage was 15mg/kg body weight. There were no refusals and no treatment withdrawals, UDCA being an extremely well tolerated drug.

In addition to recording the sex of the patients, the following measurements were taken at the time of accrual: age, serum bilirubin, serum albumin, prothrombin time, Mayo risk score, presence of oedema and positivity for AMA , ANA, ASMA, HBsAg (12 missing out of 104) and anti-HCV(12 missing out of 104). Biopsies were taken at approximately the same time. Histological staging was performed according to Ludwig et al (1978) in 103 out of the 104 patients. In the present study, patients with stage III or stage IV were classified as having ‘advanced disease’ and were compared to those with stages I or II. The reason why stage III patients were included in the present categorization of ‘advanced disease’ is that the majority of stage III patients progress to stage IV (and as no repeat biopsies were taken, it is not possible to ascertain this). Also, it is not easy to distinguish between the two stages, due to the histological heterogeneity of the disease (PJ Scheuer, 1967, Ludwig et al, 1978, cited in Jones et al, 1997). Missing values are assumed to be independent of end-point (i.e. death). Classifications of ‘II-III’ were allocated to the ‘III-IV’ group. A 3-level

oedema score was calculated as follows: 0, minimal oedema (i.e. either no oedema or oedema not requiring diuretic therapy), 0.5 moderate oedema (oedema that subsided after treatment with diuretic agents or for which no therapy was prescribed) and 1, severe oedema (oedema that persisted despite treatment with diuretic agents). This oedema scoring system is identical to that used in the Mayo model.

Log transformations were applied to the quantitative serum measurements prior to Cox regression analysis. The variables considered for inclusion in the models fitted were: Mayo risk score, age at diagnosis, log (serum bilirubin), log (serum albumin), log (prothrombin time), oedema (minimal versus either moderate or severe), stage (III-IV versus I-II) and AIC status (absent/present). Backwards and forwards selection procedures were used for model selection with entry criterion $p<0.05$ and removal criterion $p>0.05$ for each variable at each step (Wald test). Likelihood ratio tests were applied for overall model-fitting. Graphical and statistical comparisons were made with a simulated control group of PBC patients, using the updated Mayo model (see **Section 2.1.5.** for theoretical details).

It was believed that it may be more appropriate to distinguish between those PBC patients with autoimmune cholangitis (AIC) and those without AIC in making comparisons between UDCA-treated and control group subjects. Nineteen of the patients were diagnosed as having autoimmune cholangitis (AIC). Summaries of prognostic variables by AIC status are given in **Table 3.2.1.** below. When differences in the average levels of prognostic variables were assessed, the only variable for which there was weak evidence of a difference between the two groups was the serum bilirubin concentration, with the AIC group having median 1.00 mg/dL as compared to 0.80 mg/dL for the non-AIC group (Mann-Whitney test $U=547$, $p=0.028$). No differences were found between groups for the discrete variables. There was no statistical evidence provided of a difference in average risk scores between the two groups. Comparisons with simulated control groups were undertaken using the Mayo predictions for the AIC patients and non-AIC patients separately.

Table 3.2.1. Characteristics of PBC patients at baseline by autoimmune cholangitis (AIC) status: 19 AIC patients and 85 non-AIC patients

	AIC patients (n=19)	**Non-AIC patients (n=85)**
Demographic variables		
Median age in years (mean , s.e)	60 (58, 2.58)	61 (59, 1.18)
Sex, male	1 (5%)	9 (11%)
Clinical variable-Oedema	**N (%)**	
Minimal oedema	16 (84%)	77 (91%)
Moderate oedema	1 (5%)	4 (5%)
Severe oedema	2 (11%)	4 (5%)
Histologic variable– Ludwig stage[2]	**N (%)**	
Stage 1 or stage 2	10 (56%)	49 (58%)
Stage 3 or stage 4	8 (44%)	36 (42%)
Biochemical variables[3]	**Median (mean, SE)**	
Total serum bilirubin (mg/dL)	1.00 (1.18, 0.11)	0.80 (1.18, 0.17)
Serum albumin (gm/dL)	4.1 (4.2, 0.18)	4.1 (4.1, 0.06)
Prothrombin time (seconds)	13.0 (13.5, 0.35)	12.8 (13.0, 0.14)
Risk score	**Median (mean, SE)**	
Mayo model risk score	4.87 (5.03, 0.26)	4.61 (4.79, 0.13)

3.3. Patients with HCC

3.3.1 Untreated HCC patients

Between January 1992 and January 1996, 73 patients were hospitalized with HCC in the Gastroenterology Clinic. The prognostic variables available at diagnosis were: age, sex, place of residence, tumour size, Okuda stage (I, II or III), cirrhosis (present/absent), the presence of hepatitis markers (HCV, HBsAg, HCV+HBsAg, anti-Hbe and anti-HBs), α-fetoprotein AFP (ng/ml), albumin (g/L), bilirubin (mg/dL) and prothrombin time (s). The characteristics of the patients at diagnosis are presented in **Table 3.3.1.1.** below, in which it can be seen the vast majority of these patients were cirrhotics (62 out of 73, 85%) and male (61 out of 73, 84%). The missing values are assumed to be MCAR (missing completely at random, i.e. the missing data mechanism is assumed to be independent of the variables measured, pg 14, Little & Rubin, 1990).

Of these HCC patients, 48 received no therapeutic intervention. The patients who did not receive any treatment were similar with regard to prognostic factors to those who received some form of treatment e.g. the Okuda indices were for 9% I, 43% II and 48% III in the untreated patient group and 8%, 42% and 50% for I, II and III respectively in the remaining patients. Also, there were 7 females (15%) and 41 (85%) males in this subgroup whilst the other subjects consisted of 5 (20%) females and 20 males (80%). Survival analysis of the 48 untreated patients was undertaken, using December 1996 as the cut-off date for the analysis. Survival times were recorded in months. The single end-point considered was death due to HCC. There were 41 deaths during this time period, all the deaths of patients in the study being due to HCC. The median follow-up time was 5 months, the range being 1 to 33 months.

A scatter diagram (**Figure 3.3.1.1.**) indicated that there were no strong correlations between the continuous variables, although there was evidence of weak positive correlations between bilirubin concentration and prothrombin time and also between age and albumin concentration. With regard to the hepatitis markers, there was strong evidence of a negative association between HCV status and HBsAg ($p=0.0002$, chi-square test) and evidence of a positive association between anti-HBc and HBsAg

(p=0.0051, Fisher's exact test). Histograms of the variables indicated that the AFP and bilirubin concentrations had positively skewed distributions whilst the distributions of the other continuous variables appeared approximately normal. Natural logarithms were taken of the AFP and bilirubin concentrations prior to statistical analysis, resulting in approximate normality.

Table 3.3.1.1. Characteristics at diagnosis of 73 HCC patients

Characteristics		**No. (%)**
Sex	Male	61 (84)
	Female	12 (16)
Age (1)*	52-68 years	33 (46)
	69-84 years	39 (54)
Place of residence (1)	Hεraklion	28 (39)
	Rethymnon	24 (33)
	Lassithi	9 (12.5)
	Hania	9 (12.5)
	Other	2 (3)
Tumour (6)	Small	4 (6)
	Medium	18 (27)
	Large	22 (33)
	Multiple	23 (34)
Okuda index (7)	I	8 (9)
	II	28 (42)
	III	32 (49)
Cirrhosis (0)	Present	62 (85)
	Absent	11 (15)
Ascites (5)	Present	23 (34)
	Absent	45 (66)
Hepatitis indicators:		
HCV (3)	Positive	38 (54)
HBsAg (3)	Positive	18 (26)
AgHBe (25)	Positive	3 (6)
anti-HBc (25)	Positive	14 (29)
anti-HBs (25)	Positive	4 (8)
Concentrations:		*Range*
mean **AFP (ng/ml)** + SE (2)	1050 + 229.6	4 to 8650
mean **albumin (g/L)** + SE (6)	33 + 0.6	23 to 41
mean **bilirubin (mg/dL)** + SE (7)	4.7 + 0.77	0.5 to 29.4
mean **prothrombin time (s)** + SE (11)	14 + 0.1	12 to 17

*The numbers in brackets indicate the no. of missing values

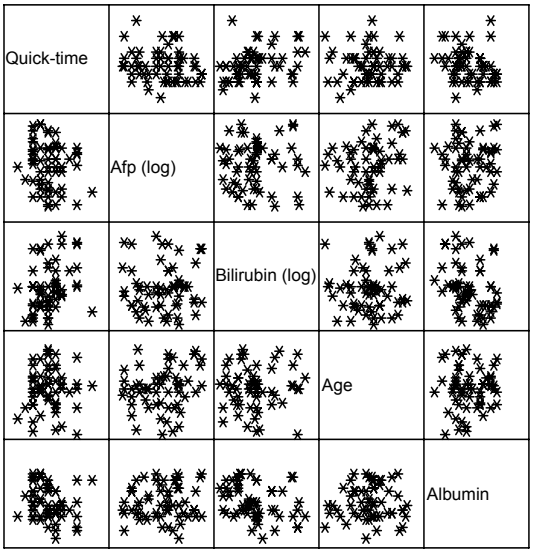


Figure 3.3.1.1. Pairwise scatterplots for the continuous prognostic variables measured at diagnosis for 48 untreated HCC patients: prothrombin time (Quick-time), AFP, bilirubin, age and albumin.

3.3.2 HCC patients treated with octreotide versus controls

A randomised, controlled study of 58 patients with HCC was undertaken, with patient accrual beginning in June 1991 and continuing until December 1995. The cut-off date of the trial was March 1996. Inclusion criteria were liver biopsy diagnosis of HCC and/or levels of alpha-fetoprotein (AFP) over 500 ng/l with compatible liver ultrasound, computed tomography scan or hepatic angiography. Exclusion criteria were small tumours judged to be suitable for surgery, variceal bleeding and hepatic encephalopathy during the previous 30 days.

The patients included in the study were randomised into one of two groups, using random number tables to determine group allocation. One group was administered 500 μg of octreotide subcutaneously in two divided doses. The other group received no treatment and served as the control group. All patients had a monthly follow-up with routine liver biochemical tests. Every two months, AFP concentrations were determined and a liver ultrasound was performed every three months. Survival times were recorded in months. Twenty-eight patients were randomised to the treatment group and 30 patients formed the controls. The end-point was death due to HCC. There were four patient withdrawals from the treatment group (none from the control group). Analysis was by intention-to-treat. There were 56 deaths during the course of the trial and 2 censored observations (as 2 subjects were withdrawn alive at the end of the trial). Follow-up ranged from 1 to 42 months.

At time of entry to the trial, the *size* of the tumour was recorded as small (4 cases), medium (9 cases), large (18 cases) or multiple (18 cases). Nine other prognostic variables were measured at entry: *AFP* concentration (ng/ml), *age*, *sex*, serum *albumin* concentration (g/L), serum *bilirubin* concentration (mg/dL), *cirrhosis* (present/absent), *place of residence*, *hepatitis* (HbsAg present, anti-HCV present, both markers present, neither marker present) and *treatment* (treated/controls). As the distributions of *AFP* concentration values and *bilirubin* concentrations were highly positively skewed, logarithms were taken of these values before any analyses were performed. The resulting variables had approximately normal distributions.

Table 3.3.2.1. Clinical and laboratory data summaries for HCC patients: 28 octreotide-treated patients and 30 untreated controls

	Patient group	
	Treated (n=28)	**Controls (n=30)**
Age in years, median (range)	69 (53 to 84)	68 (52 to 87)
Sex		
Male	23 (82%)	25 (83%)
Female	5 (18%)	5 (17%)
Cirrhosis		
Present	24 (86%)	23 (77%)
Absent	4 (14%)	7 (23%)
Mean concentration of **serum bilirubin** (mg%) (range)	5.8 (0.6 to 17.0)	6.6 (1.0 to 21.0)
Mean concentration of **serum albumin** (g/L) (range)	33.3 (25 to 41)	31.4 (25 to 40)
Child-Pugh index		
A	1 (4%)	2 (6%)
B	10 (42%)	12 (38%)
C	13 (54%)	16 (56%)
Viral markers		
HBsAg	6 (21%)	8 (27%)
Anti-HCV	15 (54%)	16 (53%)
HBsAg και anti-HCV	1 (4%)	1 (3%)
Absent	6 (21%)	5 (17%)
Tumour size		
Small (<3 cm)	3 (11%)	1 (5%)
Medium (3-8 cm)	5 (18%)	6 (19%)
Large (>8 cm)	11 (39%)	10 (33%)
Multiple	9 (32%)	13 (43%)
AFP (ng per ml)		
<100	10 (36%)	10 (33%)
100 –299	4 (14%)	5 (17%)
300 – 500	4 (14%)	5 (17%)
>500	10 (36%)	10 (33%)
Okuda stage		
I	2 (7%)	3 (10%)
II	13 (46%)	10 (33%)
III	13 (46%)	17 (57%)

The continuous variables *age*, *AFP, bilirubin*and *albumin* were grouped into 3-level factors, with approximately equal numbers at each level, for appropriate Kaplan-Meier curves to be plotted and log-rank tests performed. Two prognostic classification factors were also available at the time of entry to the study: the Child-Pugh index (A,B,C) and the Okuda stage (I,II,III). As can be seen in **Table 3.3.2.1** above, there were no major dissimilarities between the two patient groups at time of entry to the trial.

3.3.3. HCC patients treated with long-acting stomatostatin analogues versus historical controls

Following the success of the subcutaneous short-acting octreotide treatment, both in patients from the clinic (see **Section 4.1.5.1** for these results) and in other studies, it was decided in 1997 that no inoperable HCC patient would be left untreated, unless they so desired. It was decided that all patients diagnosed at the clinic between 1st October 1997 and 31st August 2000 would be accrued and treated with a long acting somatostatin analogue, with random assignment to either lanreotide (given twice every month) or octreotide LAR (given once every month). Patients in the study received no other treatment. The choice of long-acting forms as opposed to short-acting treatment was made by clinicians for patient's convenience and also as there may be a possible pharmacological advantage to long-acting forms (stable drug levels in the blood for a long period of time). There were 32 patients diagnosed with inoperable HCC during this time period, all of whom took up the option of treatment. The patients were followed up until 1st December 2000 (the cut-off date for the analysis). In order to enable comparisons of survival times with those of untreated patients, a historical control group was formed, based on medical records of HCC diagnoses at the University Hospital between 1992 and 1996.

Prognostic variables available for both groups at diagnosis included Okuda and Child-Pugh staging at diagnosis. Only 2 of the treated patients had tumours classified as Okuda III (with Child-Pugh stages B and C) and two as Child-Pugh C (Okuda II and III). In order to have homogeneity as far as possible between the treated and untreated patients with respect to the prognostic variables, the following selection restrictions were applied to the potential control group: only untreated patients with tumours of Okuda stages I or II were included and subsequently, of those selected patients with an unfavourable Child-Pugh index (stage C), only those of Okuda stage I were kept in the control group. Under these conditions, it is believed that any prior bias in favour of the treated patients (known to occur often with historical controls) was minimized, as patients who have tumours classed as Okuda stage III at the time of diagnosis are known to have a worse prognosis than those with tumours classed as I or II and similarly (although not to such a great extent) for those classed as Child-Pugh stage C. The control group thus consisted of 20 untreated historical patients.

The available prognostic variables were: *treatment* (treatment/no treatment), *age*, *sex*, *Child-Pugh index* (grades A, B or C), *Okuda stage* (I,II or III), *cirrhosis* (present/absent), *viral markers* (B,C, negative) and *typether* (the therapy administered came from one of two pharmaceutical companies; octreotide LAR 30mg vs lanreotide 30mg).[1] All variables were measured at the time of diagnosis. A summary of the prognostic variables for each group is provided in **Table 3.3.3.1.** below. The mean age of the patients at diagnosis was 69.8 (s.d. 8.2) years; the mean ages of treated and untreated patients were 70.3 (sd 8.9) years and 69.0 (sd 7.1) years respectively. Mean imputation was used for the single untreated patient of unknown age. There were no other missing values. Survival times were recorded in months, the end-point being death due to HCC. All deaths were due to HCC (i.e. there were no deaths due to other causes). There were 36 deaths and 16 censored observations, 15 of the censored observations occurring in the treated patient group and being due to termination of the study. In the control group, the censoring was due to one patient being lost to follow-up after 25 months. There were no treatment withdrawals during the study. The median follow-up time (using the 'reverse' Kaplan-Meier method) was 24 months, the ranging from 2 to 33 months.

[1] *Typether*- this variable distinguishes between the pharmaceutical companies providing the treatment, although the 2 treatments were considered equivalent. The allocation was random.

Table 3.3.3.1. Characteristics of 32 long-acting somatostatin treated HCC patients and 20 untreated historical control HCC patients[1].

	Treated patients (n=32)	**Historical controls (n=20)**
Sex		
Male	26 (81%)	19 (95%)
Female	6 (19%)	1 (5%)
Okuda stage		
I	10 (31%)	7 (35%)
II	20 (63%)	13 (65%)
III	2 (6%)	0
Child-Pugh index		
A	24 (75%)	8 (40%)
B	6 (19%)	12 (60%)
C	2 (6%)	0
Cirrhosis		
present	26 (81%)	19 (95%)
absent	6 (19%)	1 (5%)
BCLC[2]		
A_1 or A_2	11 (34%)	-
B	6 (19%)	-
C	15 (47%)	-
Viral markers		
HCV+	9 (29%)	7 (35%)
HBV+	9 (29%)	8 (40%)
Negative	13 (42%)	5 (25%)
Therapy label[2]		
Octreotide LAR	16 (50%)	-
Lanreotide	16 (50%)	-

[1]Chi-squared tests were applied to assess evidence of associations between patient group and the prognostic factors (with the exception of 'cirrhosis', for which Fisher's exact test was applied due to the small numbers in each cell). No evidence of an association between prognostic factors and treatment status was found i.e. Bonferroni-adjusted $p>0.05$ for all tests. Okuda indices II and III were merged prior to testing, as were Child-Pugh indices B and C.

[2]Not available for the historical control group

3.4. Patients with ascites

The patients suffering from ascites whose data are analysed in the present study were admitted to the Department of Internal Medicine of the University Hospital between November 1995 and February 1997. All patients were in a clinically stable condition. Two separate data sets were analysed. The first data set consisted of measurements on 50 ascites patients (**group A**), of whom 23 patients were cirrhotics whilst the remaining 27 patients had malignant peritoneal effusions. The age ranges were 45 to 83 years (median 66 years) and 39 to 88 years (median 59 years) for the patients with cirrhosis and malignant neoplasms respectively. In **group A** patients with malignant ascites, the malignancies were of the following types: ovarian carcinoma (12 patients), hepatoma (3 patients), pancreatic carcinoma (3 patients), breast carcinoma with hepatic metastases (2 patients), colon carcinoma (2 patients), peritoneal neoplasm (2 patients), stomach carcinoma (2 patients), gallbladder carcinoma (1 patient). The 23 patients with cirrhotic ascites had four types of underlying liver disease: hepatitis B (12 patients), hepatitis C (5 patients), alcoholic cirrhosis (5 patients) and primary biliary cirrhosis (1 patient). The second data set consisted of measurements on 61 subjects without signs of sepsis (**group B**, 27 males, 34 females). This group consisted of 23 cases of malignant ascites (MA), 13 cases of non-malignant ascitic exudates (NMA) and 25 cases of (non-malignant) ascitic transudates (TA)[2]. The ages of the patients in **group B** ranged from 27 to 88 years (median 62 years).

In analysing the first data set described above, the aim was to evaluate the effectiveness of the ascitic fluid to serum ratio of various acute phase proteins, immunoglobulins and cytokines complemented by other biochemical measurements such as total protein, albumin and lactate, in differentiating between malignant ascites and ascites caused by cirrhosis. In undertaking the statistical analyses it is, of course, assumed that the diagnoses of malignant and cirrhotic ascites were made independently of the measurements included in the models. Spearman's rank correlation coefficients were calculated to assess the degree of association between the biochemical ratio parameters. For **group A** patients, the non-parametric Mann-Whitney test was used to compare the levels of each biochemical ratio between the

[2] Transudates are non-malignant

two disease groups for the variables displaying a skewed distribution. The univariate tests were adjusted for multiple comparisons using the Bonferroni correction, resulting in a p-value less than 0.0019 representing significance at the 5% level (pg 384, JA Rice, 1988). Subsequently, a multivariable recursive partitioning approach was taken in order to determine the most significant biochemical predictors of the two disease groups when considering the variables simultaneously, as described in **Section 2.2.1**.

The aim in the statistical analysis of the second data set was to accurately identify the nature of a peritoneal effusion by investigating a wide array of acute-phase proteins and cytokines, complemented by other biochemical parameters, in the serum and ascitic fluid of patients. Initially, nonparametric Spearman rank correlation coefficients were calculated to assess the degree of pairwise associations between the biochemical variables. Log transformations were applied to the variables where necessary, to achieve approximate normality. For **group B** patients, one-way ANOVAs were subsequently performed to assess possible evidence of differences in the mean levels of biochemical measurements between the three groups. An indication of between which pairs of effusion types differences may lie was given by the use of Student-Newman-Keuls (SNK) contrasts. The Bonferroni correction factor was applied to adjust for multiple comparisons (pg 384, JA Rice, 1988). Multivariate discriminant analysis techniques were subsequently employed (as described in **Section 2.2.2**). Only the variables achieving significance at the univariate level were considered in the multivariate analyses.

3.5 Epidemiological survey

The seroprevalence of the viral markers HBsAg and anti-HCV in Crete were estimated in three separate epidemiological surveys. The first survey involved retrospective data from 65 219 blood donors in three Cretan prefectures taken over a five year period (1992 to 1996): Heraklion (16 792 donors), Rethymnon (16 432 donors) and Hania (31 995 donors). Only 7871 (12%) donors were female. Also estimated was the exposure to HBV in Hania and Heraklion, using positivity for the hepatitis B core antigen, HbcAb as the measure of exposure. There were no repeated measurements included in the survey. The second survey involved retrospective data obtained from 46 901 high-risk hospital patients (22 779 males, 49%) of the 281,184 (138,850 male) admissions recorded over the five-year time period. 15 391 patients (33%) were from the University Hospital in Heraklion, 21 285 patients were from the General District Hospital in Hania (45%) and 10 225 patients (22%) were from the General District Hospital in Rethymnon. The criteria for inclusion were alcoholism, altered liver function tests or exposure to standard risk factors for HBV and HCV infection (family history, professional risk, major or minor surgical operations, multiple sexual contacts etc). An exclusion criterion for HBV testing was previous vaccination for hepatitis B. Further inclusion and exclusion details are given in the corresponding publication (Koulentaki et al, 2001). All patients were tested for HBsAg, but only 73% were tested for the presence of anti-HCV whilst only patients at the Heraklion and Hania hospitals were tested for HBcAb.

In each of the two above surveys, the standard large sample normal approximation to the binomial distribution was used to assess differences in prevalence between regions (pg 123-5, Armitage & Berry, 1987). In addition, Cochran's test was applied to test for overall differences between sexes, whilst accounting for different sample sizes from each region (pg 380-4, Armitage & Berry, 1987). Finally, an odds ratio approach was taken to compare risks by gender, with the estimation of 95% C.I.s for the odds ratio estimates (pg 458, Armitage & Berry, 1987, pg 21, A. Agresti, 1984).

The third survey was a community-based serosurvy. The design was, in fact, that of a two-stage stratified sampling process. For stage 1, 8 regions in Heraklion and 5 rural areas (villages) were selected at random with probabilities proportional to population size. Stage 2 involved samples being drawn at random from each of the stage 1 units,

using uniform sampling fractions. In total, 446 (50 % female) and 479 (40% female) subjects were tested in urban and rural areas respectively. Significance tests and confidence intervals were again obtained using standard tests for differences between proportions. It is known, however, that the simple random sampling formulae may underestimate the true standard errors, contributing more precision to the sample estimates than they actually have (pg 202, Moser & Kalton, 1971).

4.1 Survival analysis

4.1.1 The natural history of cirrhosis

A. Prognosis for compensated cirrhotics

The Kaplan-Meier PL estimate of the 'decompensation-free' function is given in **Figure 4.1.1.1**. The median overall time to decompensation was 58 months with 95% confidence limits (CL) of 51 and 65 months. The two prognostic factors available at diagnosis that were found to have a significant effect on the time to decompensation (using the log rank test and univariate Cox PH models) were the aetiology of the cirrhosis and the sex of the patient (**Table 4.1.1.1**).

Table 4.1.1.1. Estimated median time to decompensation and log rank test results by prognostic factors sex, age and aetiology (n=306[§])

Factor	n	Time to decompensation			Logrank statistic (df)	p value
		Median	Lower 95% CL	Upper 95% CL		
Overall: 150 events, 156 (51%) censored	**306**	**58**	**51**	**65**		
Sex					9.20 (1)	0.0024
Male	168	50	38	62		
Female	138	72	53	91		
Age at diagnosis (12 missing)					1.30 (1)	NS
<64 yrs	144	66	48	83		
>=64 years	162	57	46	68		
Type of cirrhosis					30.50 (4)	<0.0001
Alcohol	56	35	19	51		
Alcohol+virus*	17	31	16	46		
Hepatitis B	45	36	20	52		
Hepatitis C	145	81	45	117		
Cryptogenic/Other	43	58	55	61		

[§]306 of the 312 patients were included in the survival analysis (5 had missing dates of diagnosis and 1 had missing date of decompensation).
* 6 had alcohol+HBV, 11 had alcohol+HCV

Kaplan-Meier estimates of the proportion of patients remaining compensated (i.e. 'decompensation-free') against time to decompensation according to cirrhosis aetiology are depicted in **Figure 4.1.1.2**. It is clearly seen that those with hepatitis C have longer times to decompensation than the other groups (overall log rank test statistic 30.5 on 4 df, $p<0.0001$, risk of decompensation for those with hepatitis C being 0.57 times that of cryptogenics, with 95% CL 0.34 and 0.96; the median time to decompensation for hepatitis C cirrhotics was 81 months with 95% CL of 45 and 117 months.

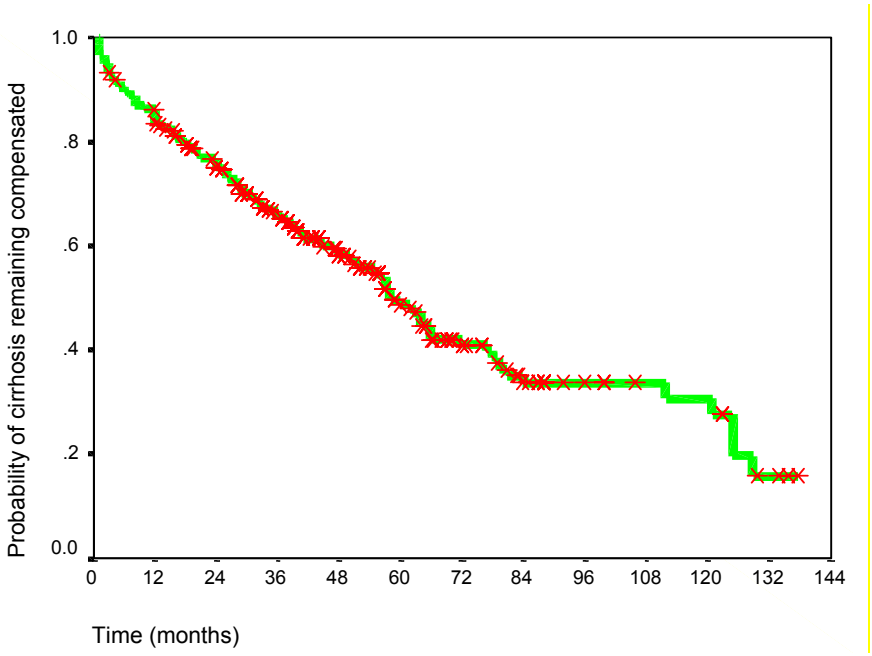


Figure 4.1.1.1. Kaplan-Meier estimate of the 'decompensation-free' function for the 306 patients diagnosed with compensated cirrhosis.

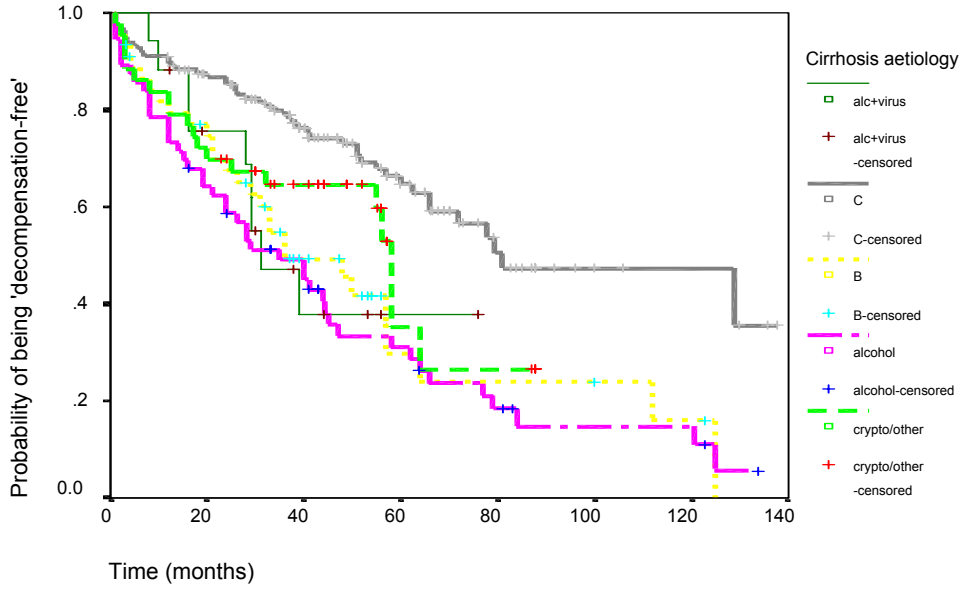


Figure 4.1.1.2. Kaplan-Meier estimates of the 'decompensation-free' functions for patients diagnosed with compensated cirrhosis by aetiological group.

The 3-, 5- and 7- year decompensation rates are presented by cirrhosis aetiology in **Table 4.1.1.2,** from which it can be seen that 65% of the cirrhotics initially diagnosed as compensated remain in the same state after 3 years, with this percentage falling to 49% after 5 years and 34% after 7 years. For those patients who are anti-HCV

positive, the decompensation-free percentage after 7 years is 47% (95%C.I. 34% to 61%) as contrasted with 15% for those with alcohol as aetiology (with 95% C.I. 3% to 26%).

Table 4.1.1.2. Estimated percentage of cirrhosis patients remaining decompensation-free 3, 5 and 7 years after diagnosis by aetiology and gender (n=306)

	No. of patients	Cumulative percentage of decompensation-free patients								
		3 y			5 y			7 y		
		%	Lower 95% CL	Upper 95% CL	%	Lower 95% CL	Upper 95% CL	%	Lower 95% CL	Upper 95% CL
All	306	65	60	71	49	42	56	34	26	42
Aetiology Alcohol	56	49	36	62	29	16	42	15	3	26
Alcohol+virus*	17	47	21	73	38	11	64			
HBsAg positive	45	49	34	65	30	12	48	24	6	42
anti-HCV posititve	145	79	72	86	65	55	74	47	34	61
No aetiology given/cryptogenic	43	65	50	79	35	12	59	27	3	50
Sex male	168	60	52	68	44	35	53	25	14	36
female	138	72	64	80	54	44	64	43	32	55

*The maximum follow-up time was 76 months

Table 4.1.1.3. Estimated relative decompensation rates by significant prognostic factors after fitting a Cox PH model for time to decompensation (n=299*)

Factor	n	Relative rate	Lower 95% CL	Upper 95% CL	Wald statistic	p value
Age at diagnosis	299	1.02	1.00	1.03	4.49 (1)	0.034
Type of cirrhosis					29.81 (4)	<0.0005
Alcohol	55	1.72	1.00	2.97		
Alcohol+virus	16	1.57	0.67	3.66		
Hepatitis B	45	1.48	0.83	2.67		
Hepatitis C	141	0.58	0.34	1.00		
Cryptogenic/No aetiology given	42	1				

*Initially 312 patients, but 13 patients had unknown values for at least 1 variable, including 6 with unknown time to decompensation

The results of the multivariate Cox PH model are presented in **Table 4.1.1.3** above, in which it be seen that age at diagnosis and cirrhosis aetiology are both significant prognostic factors for the time to decompensation. A year's increase in the age at diagnosis leads to a 2% increase in the hazard of decompensation. The risk of decompensation of cirrhotics with alcohol as aetiology is estimated to be 72% higher

than that of cryptogenic cirrhotics, with 95% CL of 0% higher and 197% higher, whereas type C patients have an estimated risk 42% lower than the risk of cryptogenics, with 95% CL of a 66% decrease and equal risk. **Figure 4.1.1.3** is a plot of the martingale residuals against the age at diagnosis variable, from which it appears that age can be adequately used in the model without transformation (note that the censored cases have been distinguished from those in which decompensation occurs, the former taking negative values as described in **Section 2.1.2.2.1.**) as there is roughly random scatter about 0.

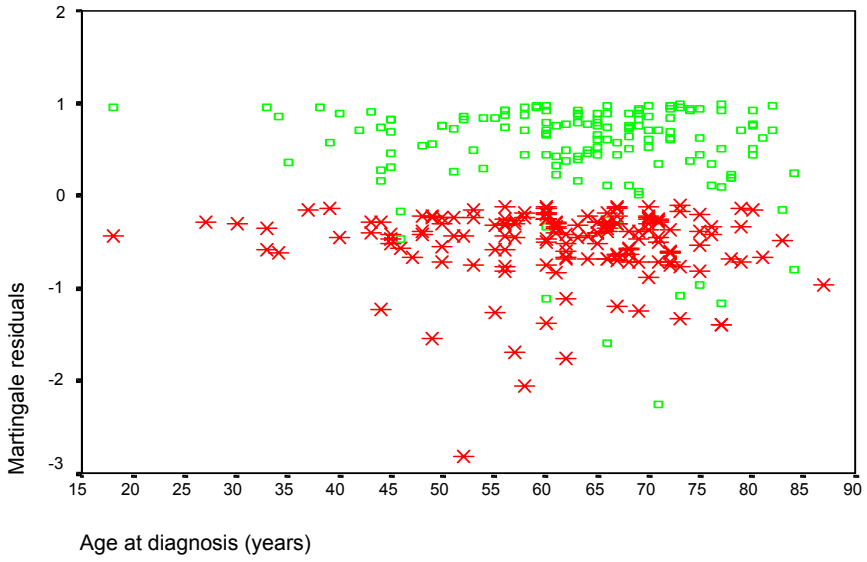


Figure 4.1.1.3. A plot of martingale residuals for the Cox time to decompensation model by age at diagnosis; stars represent decompensation-free cases and squares represent patients who decompensate during the study period.

A log-cumulative hazard plot of the Cox-Snell residuals is given in **Figure 4.1.1.4**, from which it may be inferred that the model fits the data satisfactorily, as the plot is fairly close to a straight line. The 6 smallest (most negative) Cox-Snell residuals, including the log-residual of –6.1, correspond to patients who decompensate one month after diagnosis.

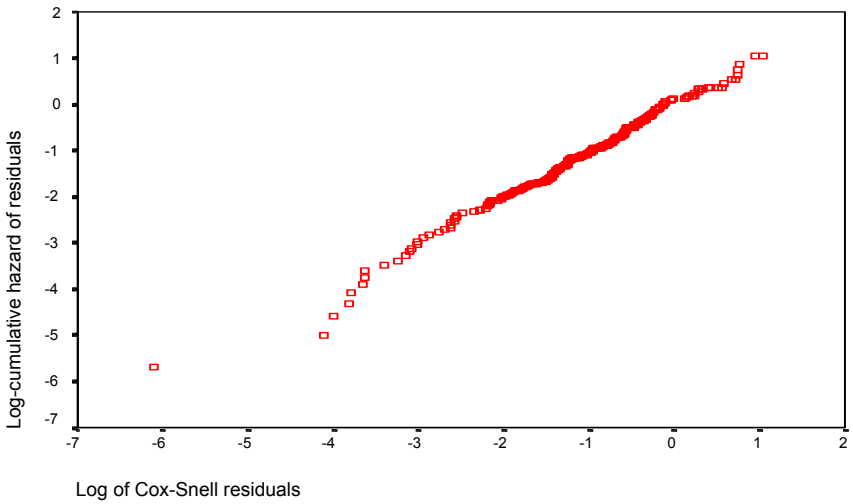


Figure 4.1.1.4. A log-cumulative hazard plot of the Cox-Snell residuals for the Cox PH model fitted for time to decompensation.

Bootstrap analysis confirmed the stability of the Cox model for time to decompensation in compensated cirrhotics. In **Table 4.1.1.4** below it can be seen that the estimated regression coefficients and their SEs prior to resampling are very similar to the bootstrap results. The median and mean of the coefficients $\hat{\beta}*$ were also in close agreement. Based on the 90% CL for both the BCa and the empirical percentiles, all coefficients except HBV (versus cryptogenic) and alcohol/virus (versus cryptogenic) were significantly different from zero, as in the original model.

Table 4.1.1.4 Bootstrap estimates of regression coefficients and standard errors based on the Cox decompensation model (1000 replicates)

Variable	**Regression Coefficient**		**Bootstrap regression coefficient**		**Estim-ated bias**	**Median & 90% empirical confidence limits**			**Median & 90% BCa confidence limits**		
	$\hat{\beta}$	$se(\hat{\beta})$	$\hat{\beta}*$	$se(\hat{\beta}*)$		**5%**	**50%**	**95%**	**5%**	**50%**	**95%**
Age	0.016	0.007	0.017	0.008	0.001	0.004	0.017	0.029	0.003	0.015	0.028
Cirrhosis aetiology:											
Alcohol	0.544	0.279	0.565	0.311	0.021	0.068	0.560	1.081	0.003	0.527	1.035
HBV	0.400	0.299	0.407	0.311	0.007	-0.122	0.402	0.919	-0.139	0.397	0.893
HCV	-0.543	0.272	-0.542	0.296	-0.001	-1.014	-0.553	-0.048	-1.001	-0.535	-0.014
Alcohol + virus	0.454	0.432	0.424	0.470	-0.030	-0.342	0.443	1.155	-0.309	0.474	1.177

Histograms of the empirical distribution of replicated regression coefficients for each variable with a smoothed density estimate are provided in **Figure 4.1.1.5**, from which

it can be seen that the distributions are very close to normal. Using jackknife after bootstrap techniques to assess the influence of each of the observations on the bias, no highly influential points were detected (the maximum absolute relative influence on the bias being less than 5 for all observations).

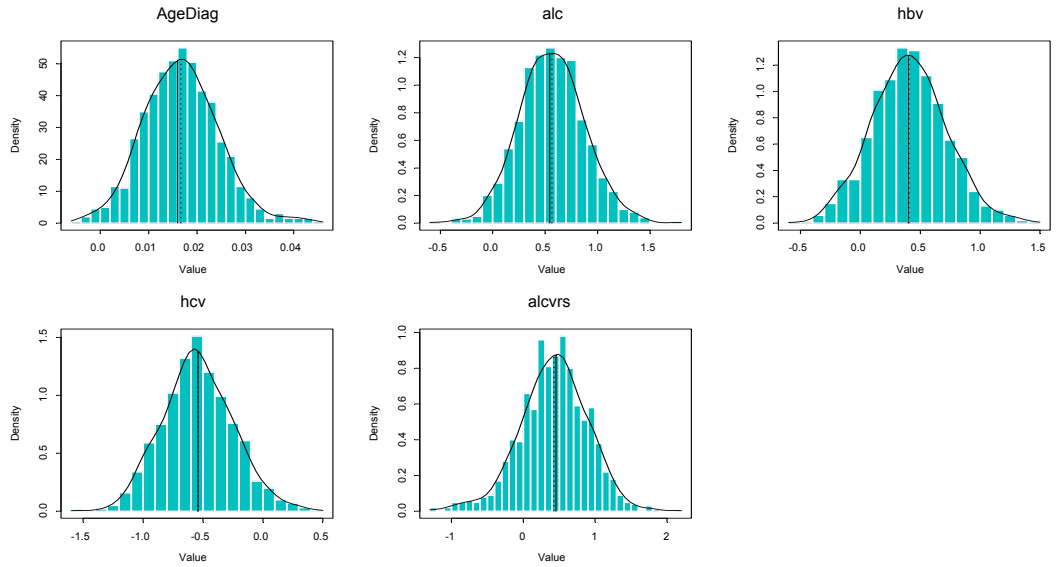


Figure 4.1.1.5. Histograms of the empirical distributions of parameter replicates for the Cox time-to-decompensation model (B=1000)

A prognostic index (**PI**) for decompensation was derived from the Cox model as :

PI = 0.016 * (age –62.29) + 0.54*Ind{alcohol} + 0.40*Ind{ HBV}- 0.54*Ind{HCV} + 0.45 *Ind{alcohol+ viral hepatitis}

where Ind{X}=1 if X is the aetiology, 0 otherwise.

If a compensated cirrhosis patient has cryptogenic cirrhosis aetiology, the PI is simply 0.016 * (age –62.29). The PI for the patients in the study ranges from -0.94 to 0.88. **Figure 4.1.1.6** shows the estimated probability of being decompensation-free at 3, 5 and 7 years as a function of the PI. The SPSS syntax for derivation of this figure is provided in **Appendix F**. For example, a man (or woman) aged 44 with HBV as the underlying cause will have PI of 0.016*(44-62.29)+0.40=0.11 and will therefore have estimated probabilities of being decompensation-free for at least 3, 5 and 7 years after diagnosis of 0.49, 0.29 and 0.15 respectively. Once the PI has been calculated, these probabilities can be estimated graphically using **Figure 4.1.1.6.** If the same person had HCV as underlying aetiology, the corresponding PI would be 0.016*(44-62.29)-

0.54 =-0.83 with estimated probabilities of being decompensation-free for at least 3, 5 and 7 years after diagnosis of 0.75, 0.62 and 0.48 respectively.

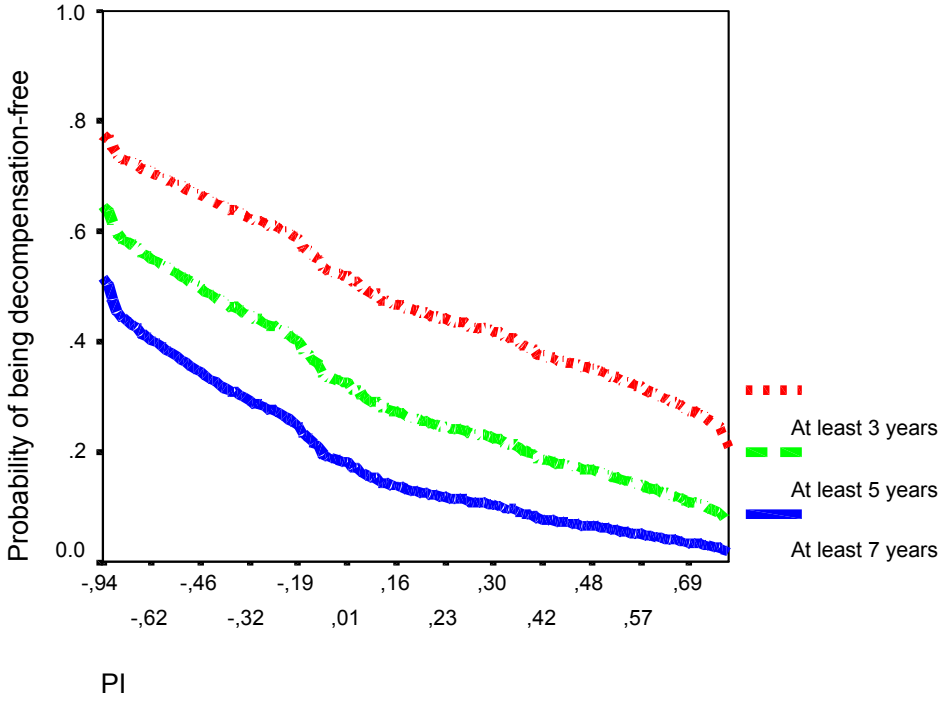


Figure 4.1.1.6 Estimated probability of a compensated cirrhosis patient diagnosed at the Clinic remaining decompensation-free for at least 3, 5 and 7 years after diagnosis, as a function of the prognostic index (PI).

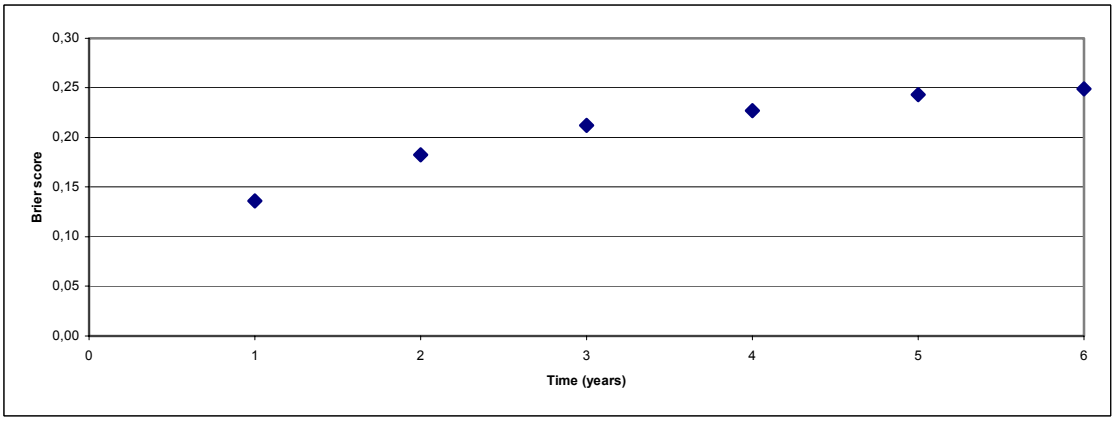


Figure 4.1.1.7. The Brier score for the PI derived from the Cox model for time to decompensation in 306 patients diagnosed with compensated cirrhosis.

Figure 4.1.1.7 is a graph of the Brier score for the data as a function of time. For a short time after diagnosis, it can be seen that the Cox model predictions are accurate,

as the Brier score is relatively low (e.g. 0.14 at 1 year). From 5 years onwards, however, the score approaches 0.25, 0.25 being the score when the trivial prediction $\hat{\pi}(t^*) = 0.5$ is made for all patients. $\mathbf{R^2}$ takes values of 1% and 2% at 1 and 2 years respectively and reaches a maximum of 6% after 3 years, but from 5 years on, there is no advantage of the Cox model over the simple Kaplan-Meier estimate (as the R^2 value falls to 4% after 4 years and 0% after 5 years). This may be expected, as the median follow-up time is just under 5 years and at this time the Kaplan-Meier estimate approaches 0.5, a situation where predictions are harder to make than initially when almost all patients are alive (Graf et al, 1999).

The median survival time for patients presenting with compensated cirrhosis was 10.5 years with 95% C.I. 103 to 149 months. The Kaplan-Meier PL estimate of the overall survival function is given in **Figure 4.1.1.8** below. The two prognostic factors available at diagnosis that were found to have a significant effect on the survival time were again the aetiology of the cirrhosis and the sex of the patient (**Table 4.1.1.5** below).

Table 4.1.1.5. Estimated median survival times and log rank test results for patients diagnosed as having compensated cirrhosis by sex, age and aetiology (n=307*).

	n	Survival time in months for compensated cirrhotics			Logrank statistic (df)	p value
		Median	Lower 95% CL	Upper 95% CL		
Overall: 70 events, 237 (77.2%) censored	**307**	**126**	**103**	**149**		
Sex (3 missing)	304				11.58 (1)	0.0007
Male	145	106	78	134		
Female	162	114**	104	123		
Age (10 missing)	297				1.52 (1)	NS
<64 yrs	145	130	91	169		
>=64 years	162	126	84	168		
Cirrhosis aetiology	304				16.54 (4)	0.0024
Alcohol	56	120	101	139		
Alcohol+virus	17	84	68	100		
Hepatitis B	45	85**	66	103		
Hepatitis C	146	113**	103	124		
Cryptogenic/Other	43	72	65	83		

*Initially 312 patients, but 5 have missing survival time

** Estimates of the mean survival time (as it was not possible to estimate the variability of the median)

The mean survival time for type C cirrhotics was estimated to be 113 months with 95% C.I. 103 to 124 months. Compensated type C cirrhosis patients appear to have

longer survival times on average than the other aetiological groups (overall log rank test statistic 16.5 on 4 df, p<0.002, the hazard for those with hepatitis C as aetiology being 66% lower than that of cryptogenics, with 95% CL of 83% lower and 31% lower. Female compensated cirrhotics have longer survival times on average, with death risk for females estimated to be 58% lower than that of males, with 95% CL 75% lower and 30% lower than the risk for males.

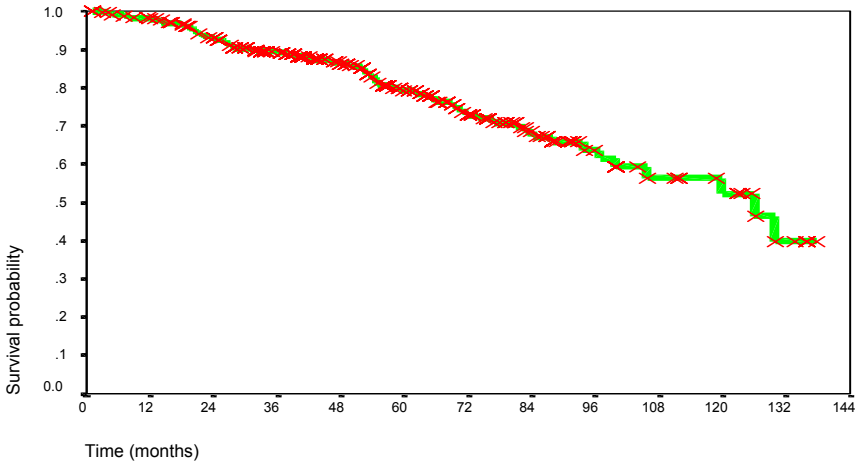


Figure 4.1.1.8. Kaplan-Meier estimate of the survivor function for the 306 patients diagnosed with compensated cirrhosis.

Kaplan-Meier PL estimates of the survivor function according to cirrhosis aetiology are depicted in **Figure 4.1.1.9.** below.

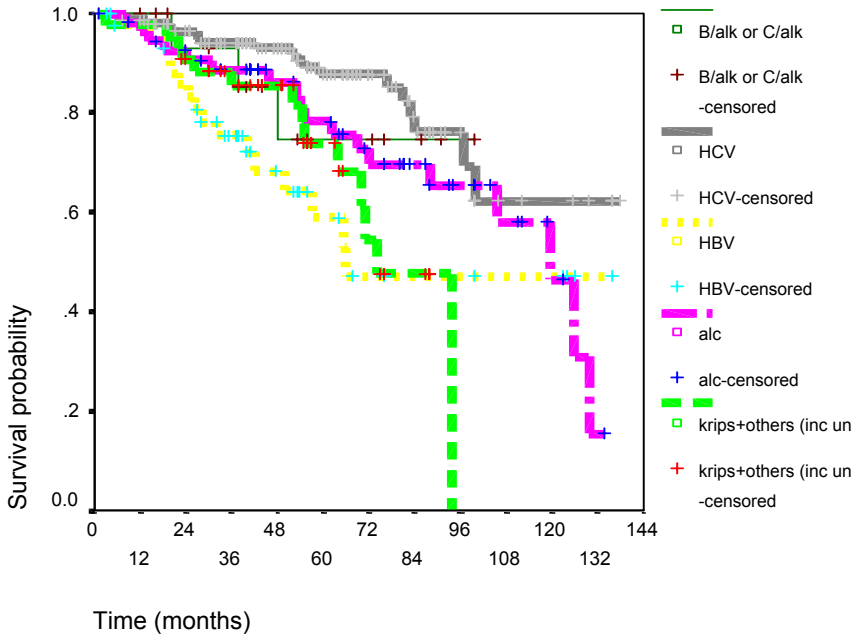


Figure 4.1.1.9. Kaplan-Meier estimates of the survivor function for the 306 patients diagnosed with compensated cirrhosis by cirrhosis aetiology.

The 3-, 5- and 7- year survival rates are presented by cirrhosis aetiology in **Table 4.1.1.6** below, from which it can be seen that 89% of the patients with an initial diagnosis of compensated cirrhosis remain alive after 3 years, with this percentage decreasing to 79% after 5 years and 67% after 7 years. Seven years after diagnosis, 76% of HCV patients and 70% of alcoholics remain alive, as contrasted with 47% of HBV patients and 48% of those with cryptogenic cirrhosis.

Table 4.1.1.6. Estimated percentage of patients surviving 3, 5 and 7 years after diagnosis of compensated cirrhosis (n=307)

	No. of patients	Survival percentages for compensated cirrhotics								
		3 y			5 y			7 y		
		%	Lower 95% CL	Upper 95% CL	%	Lower 95% CL	Upper 95% CL	%	Lower 95% CL	Upper 95% CL
All cirrhotics	307	89	85	93	79	74	85	67	59	75
Aetiology										
Alcohol	56	89	80	97	78	66	90	70	55	84
Alcohol+virus*	17	85	66	100	74	49	100			
HBsAg +	45	75	62	89	64	48	81	47	27	68
anti-HCV +	146	94	90	98	88	81	94	76	64	88
No aetiology given/cryptogenic	43	85	75	96	74	58	89	48	25	71
Sex male	168	85	80	91	72	64	80	60	50	71
female	139	94	89	98	87	80	94	76	64	87

* There were no events after 5 years or more after presentation. This may be due to the small numbers in this sub-group by this time.

The results of the multivariate time-fixed Cox PH model for overall survival prognosis of compensated cirrhosis using the three variables available at diagnosis are presented in **Table 4.1.1.7** below. Sex, age at diagnosis and cirrhosis aetiology can all be seen to be significant prognostic factors for the survival of these patients. Females have an estimated death rate of 0.37 relative to males, with 95% CL of 0.20 and 0.69. An estimate of the death rate of cirrhotics with with anti-HCV positive is 0.45 compared to cryptogenics, with 95% CL of 0.22 and 0.94. An increase of one year in age results increases the risk of death by 3%, with 95% CL of 0% (i.e. no change in risk) and 5%.

Table 4.1.1.7. Estimated relative death rates by prognostic factor after fitting a Cox PH model for patients diagnosed with compensated cirrhosis (n=297*)

Factor	*Relative rate*	Lower 95% CL	Upper 95% CL	Wald statistic	p value
Sex					
Female	0.37	0.20	0.69	9.86 (1)	0.0017
Age at diagnosis	1.03	1.00	1.05	5.15(1)	0.0233
Type of cirrhosis				10.10 (4)	0.0388
Alcohol	0.53	0.25	1.14		
Alcohol+virus	0.34	0.07	1.54		
Hepatitis B	1.07	0.51	2.27		
Hepatitis C	0.45	0.22	0.94		
Cryptogenic/Other	1				

*Initially 312 patients, but 15 patients had unknown values for at least 1 variable, including 5 with unknown survival time

The stability of the above model was investigated using bootstrap analysis. In **Table 4.1.1.8** below it can be seen that the estimated regression coefficients and their SEs prior to resampling are similar to the bootstrap results, although the bias in the alcohol/virus aetiology coefficient stands out at -0.56. The median and mean of the $\hat{\beta}*$ were in close agreement for all variables other than the alcohol/virus aetiology indicator variable (mean coefficient –1.643, empirical median –1.136, BCa median –1.046). Based on the 90% CL for both the BCa and the empirical percentiles, the coefficients for sex, age and HCV were significantly different from zero, as in the original model. Histograms of the empirical distribution of replicated regression coefficients for each variable with a smoothed density estimate are provided in **Figure 4.1.1.10.**, from which it can be seen that the distributions are close to normal for all apart from the alcohol+viral hepatitis aetiology indicator variable, for which the distribution appears to be bimodal. Using jackknife after bootstrap techniques to assess the influence of each of the observations on the bias, there were found to be two highly influential points in the distribution of the alcohol+viral hepatitis aetiology indicator variable (with absolute relative influence on the bias of 10.07 and 9.78, see **Figure 4.1.1.11**). When the Cox model was refitted omitting the influential observations, the empirical distribution histograms appeared close to normal for all variables (**Figure 4.1.1.12**) and no highly influential points were detected. The regression coefficients, and hence the relative risks, changed only very slightly for variables other than the alcohol+viral hepatitis variable (see **Table 4.1.1.9** below). In fact, only the RR for the alcohol plus virus aetiological category changed to more than the second decimal place, changing from 0.34 to 0.003.

Table 4.1.1.8 Bootstrap estimates of regression coefficients and standard errors based on the Cox survival model for compensated cirrhotics (1000 replicates)

Variable	Regression Coefficient		Bootstrap regression coefficient		Estimated bias	Median & 90% empirical confidence limits			Median & 90% BCa confidence limits		
	$\hat{\beta}$	$se(\hat{\beta})$	$\hat{\beta}*$	$se(\hat{\beta}*)$		5%	50%	95%	5%	50%	95%
Sex	-0.993	0.316	-1.010	0.344	-0.016	-1.586	-1.000	-0.450	-1.581	-0.985	-0.446
Age	0.025	0.011	0.026	0.011	0.001	0.008	0.026	0.046	0.005	0.024	0.044
Cirrhosis aetiology:											
Alcohol	-0.627	0.388	-0.631	0.401	-0.004	-1.283	-0.642	0.038	-1.248	-0.608	0.059
HBV	-0.077	0.380	0.121	0.422	0.044	-0.566	0.113	0.813	-0.723	0.046	0.685
HCV	-0.789	0.370	-0.778	0.380	-0.011	-1.397	-0.791	-0.143	-1.404	-0.791	-0.155
Alcohol + virus	-1.082	0.772	-1.643	1.875	-0.560	-6.283	-0.136	0.131	-6.126	-1.046	0.179

Table 4.1.1.9 Regression coefficients and standard errors for the Cox survival model for compensated cirrhotics, before and after the removal of observations found to be influential using jackknife after bootstrap techniques.

Variable	Cox model regression coefficients and standard errors prior to stability analysis				Cox model regression coefficients and standard errors following stability analysis			
	$\hat{\beta}$	$SE(\hat{\beta})$	Z	p	$\hat{\beta}$	$SE(\hat{\beta})$	Z	p
Sex	-0.993	0.316	-3.141	0.002	-0.988	0.317	-3.120	0.002
Age	0.025	0.011	2.270	0.023	0.023	0.011	2.102	0.036
Alcohol	-0.627	0.388	-1.615	0.110	-0.645	0.388	-1.661	0.097
HBV	0.077	0.380	0.202	0.840	0.073	0.380	0.192	0.850
HCV	-0.789	0.370	-2.131	0.033	-0.794	0.370	-2.146	0.032
Alcohol+viral hepatitis	-1.082	0.772	-1.402	0.160	-5.894	8.015	-0.735	0.460

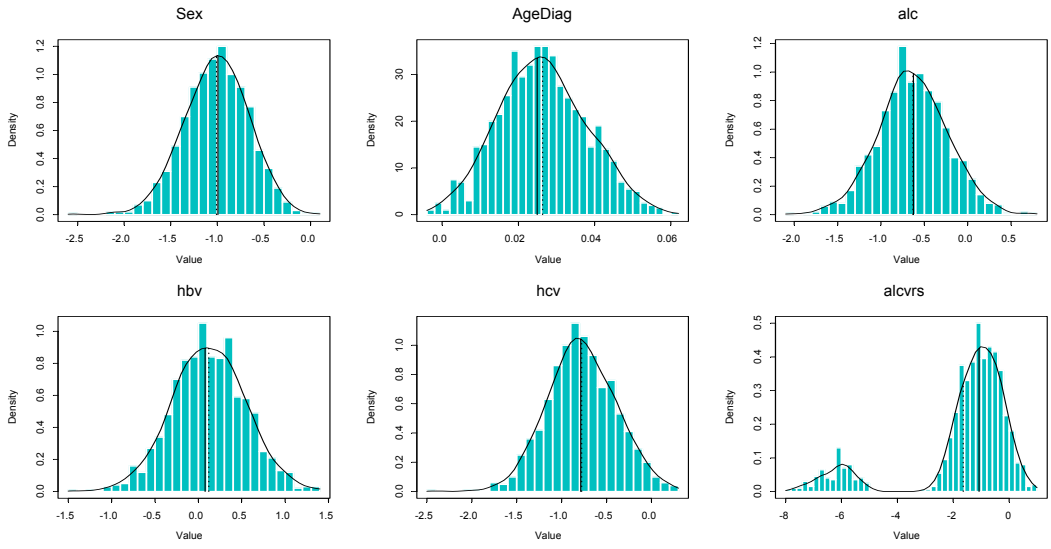


Figure 4.1.1.10. Histograms of the empirical distributions of parameter replicates for the Cox survival model for compensated cirrhotics (B=1000)

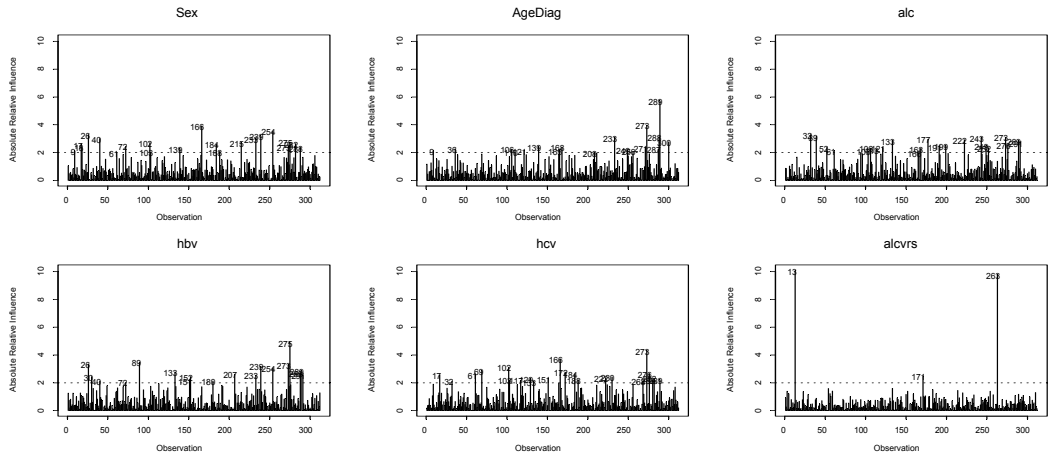


Figure 4.1.1.11. Influence plots for regression coefficients in the Cox model for survival time of compensated cirrhotics.

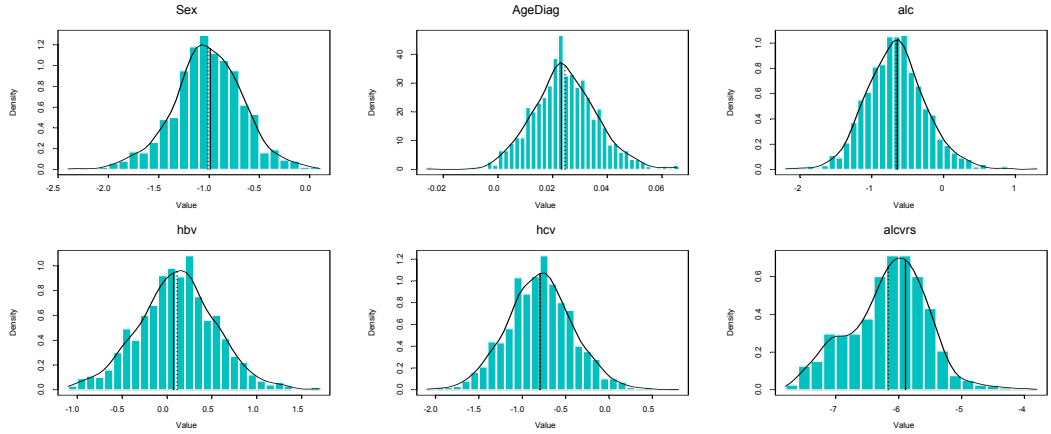


Figure 4.1.1.12. Histograms of the empirical distributions of parameter replicates for the Cox survival model for compensated cirrhotics excluding influential observations (B=1000)

It is known that HCC and decompensation may occur at some date following diagnosis of compensated cirrhosis. The model derived above can be extended to take account of the occurrence of HCC and/or decompensation. In **Table 4.1.1.10a** are displayed the results of entering the binary factor HCC (no, yes) into the prognostic model derived previously, as a time-dependent variable. In the time-dependent model, the coefficients represent the effects that the covariates have both at entry and any time thereafter (and so cannot be interpreted as easily as those in time-fixed models, and cannot be interpreted as direct effects on the hazard at entry). The presence of the HCC variable overrides the effect of sex and age at diagnosis. A model incorporating the occurrence of decompensation as a time-dependent variable, in addition to the variables sex, age at diagnosis and cirrhosis aetiology is given in **Table 4.1.1.10b.** It can be seen that following backwards selection procedures sex, cirrhosis aetiology

and decompensation (yes/no) remain as significant prognostic factors. Finally, a model is fitted in which both HCC (no, yes) and the occurrence of decompensation (no, yes) are introduced as time-dependent variables in addition to the other prognostic variables (**Table 4.1.1.10c**). The final model contains only the two time-dependent variables, showing that once the occurrence of HCC or decompensation has been accounted for, the other variables (sex, aetiology, age) are not of significance in the survival prognosis. The corresponding SPSS syntax for the time-dependent model is given in **Appendix C**.

Table 4.1.1.10a. Factors associated with survival for those diagnosed with compensated cirrhosis, including HCC as a time-dependent covariate in the prognostic model (n=297)

Factor	**Regression coefficient (β_i)**	**SE (β_i)**	**p value**
HCC	3.03	0.297	0.0000
Cirrhosis aetiology			0.0138
Alcohol	-0.39	0.376	
Alcohol+virus	-0.88	0.764	
Hepatitis B	-0.22	0.393	
Hepatitis C	-1.15	0.371	
Cryptogenic/Other			

Table 4.1.1.10b. Factors associated with survival for those diagnosed with compensated cirrhosis, including the occurrence of decompensation as a time-dependent covariate in the prognsotic model (n=297)

Factor	**Regression coefficient (β_i)**	**SE (β_i)**	**p value**
Sex			
Female	-1.14	0.355	0.0013
Decompensate	3.79	0.480	0.0000
Cirrhosis aetiology			0.004
Alcohol	-0.82	0.380	
Alcohol+virus	-1.33	0.767	
Hepatitis B	0.49	0.380	
Hepatitis C	0.11	0.415	
Cryptogenic/Other			

Table 4.1.1.10c. Factors associated with survival for those diagnosed with compensated cirrhosis, including the occurrence of decompensation and HCC as time-dependent covariates in the prognostic model, considering also sex, age and type of cirrhosis (n=297)

Factor	**Regression coefficient (β_i)**	**SE (β_i)**	**p value**
Decompensation	3.64	0.478	0.0000
HCC	2.83	0.306	0.0000

B. Prognosis for decompensated cirrhotics

The Kaplan-Meier PL estimate of the survival function for those presenting with decompensated cirrhosis is given in **Figure 4.1.1.13** below. The median overall survival time was 59 months (with 95% C.I. 43 to 76 months). The single prognostic factor available at diagnosis that were found to have a significant effect on the time to decompensation (using the log rank test and univariate Cox PH models) was the type of decompensation (**Table 4.1.1.11**).

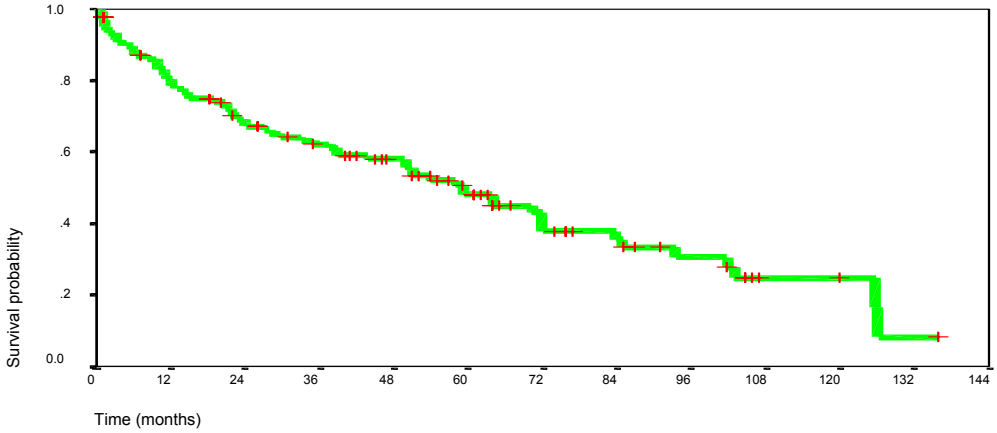


Figure 4.1.1.13. Kaplan-Meier estimate of the survival function for the 138 patients who present with decompensated cirrhosis.

Table 4.1.1.11. Estimated median survival time and log rank test results for the 138 patients who present with decompensation by prognostic factors sex, age, type of decompensation and cirrhosis aetiology

	n	Estimated survival time in months for decompensated cirrhotics			Logrank statistic (df)	p value
		Median	Lower 95% CL	Upper 95% CL		
Overall: 66 events, 72(52.2%) censored	138	59	43	76		
Sex					0.02 (1)	NS
Male	107	58	47	69		
Female	31	72	37	106		
Age					0.06 (1)	NS
<64 yrs	67	59	43	75		
>=64 years	71	58	35	81		
Type of decompensation					21.38 (3)	0.0001
Variceal bleeding	22	64	33	95		
Hepatic encephalopathy/hepatic encephalopathy+other	10	5	2	8		
Ascites	89	64	49	79		
Other/Unknown	17	54	16	92		
Cirrhosis aetiology					5.00 (3)	NS
Alcohol/Alcohol+virus	73	64	41	87		
Hepatitis B	18	14	0	34		
Hepatitis C	28	59	33	86		
Cryptogenic/Other/Missing	19	64	25	102		

Kaplan-Meier estimates of the survivor function according to the nature of the decompensation are depicted in **Figure 4.1.1.14**. below.

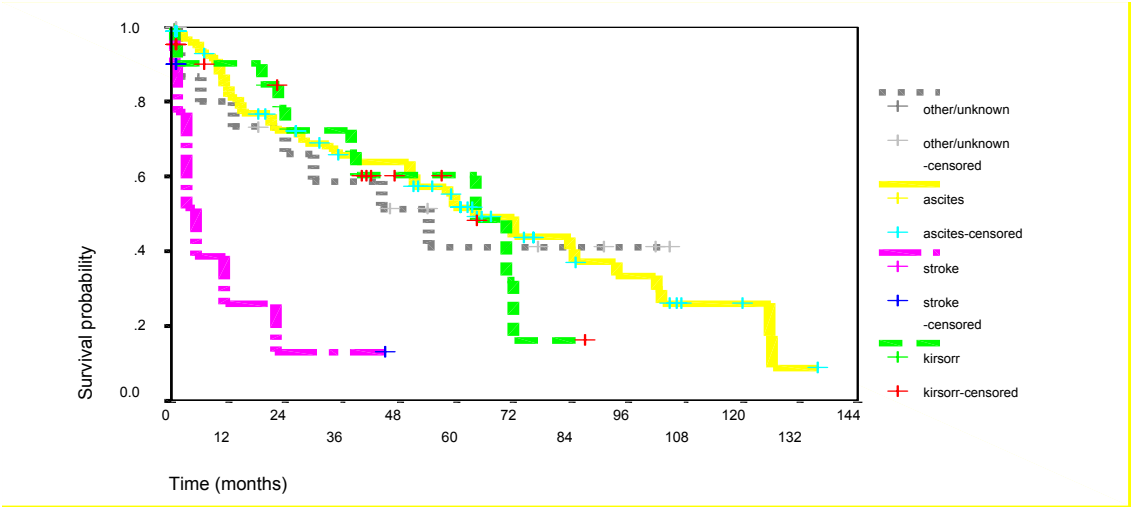


Figure 4.1.1.14. Kaplan-Meier estimate of the survival function for the 138 patients who present with decompensated cirrhosis by type of decompensation.

It is clearly seen that those suffering from hepatic encephalopathy have shorter survival times than the other groups (overall log rank test statistic 21.4 on 3 df, p=0.0001), with the hazard for those with hepatic encephalopathy being 4.95 times that of those who display variceal bleeding (95%C.I. 1.84 to 13.32). The median survival time for cirrhotics presenting with hepatic encephalopathy was 5 months with 95% C.I. 2 to 8 months.

The percentages of decompensated cirrhotics surviving after 3, 5 and 7 years are presented by type of decompensation in **Table 4.1.1.12**. Three years after presentation with decompensated cirrhosis, 62% of patients remain alive, whereas this falls to 48% after 5 years and 36 % after 7 years. These percentages vary greatly between the decompensation groups e.g. after 3 years only 13% of those who suffered from hepatic encephalopathy remain alive, as contrasted with 72% of those with variceal bleeding and 66% of those with ascites.

Table 4.1.1.12. Estimated percentage of patients surviving 3, 5 and 7 years after presentation with decompensated cirrhosis by aetiology and type of decompensation (n=138)

	No. of patients	Survival percentages for decompensated cirrhotics								
		3 y			5 y			7 y		
		%	Lower 95%CL	Upper 95%CL	%	Lower 95%CL	Upper 95%CL	%	Lower 95%CL	Upper 95%CL
All cirrhotics	138	62	53	71	48	38	58	36	26	45
Cirrhosis aetiology										
Alcohol/Alcohol+virus	73	66	54	79	52	38	66	41	25	57
Hepatitis B	18	34	9	60	26	2	50	13	0	34
Hepatitis C	28	69	49	88	50	28	71	30	8	52
Cryptogenic/Other/Missing*	19	64	40	87	55	29	81			
Type of decompensation										
Variceal bleeding	22	72	52	93	60	37	83	32	0	64
Hepatic encephalopathy**	10	13	0	36						
Ascites	89	66	54	77	52	39	64	40	27	54
Other/unknown***	17	59	33	84						

*The last death was at 5 years, 4 months. After this there were only 3 censored observations
**Or hepatic encephalopathy plus other complication: the longest survival time here was 3 years, 9 months
***The last death was at 4 years, 6 months. After this there were only 3 censored observations

The multivariate Cox PH model fitted stepwise selection procedures indicated that age was also a significant prognostic factor, when considered jointly with the type of decomposition (**Table 4.1.1.13**).

Table 4.1.1.13. Cox estimates of relative death rates by significant prognostic factor for those who present with decompensation, initially considering sex, age, type of decompensation and type of cirrhosis (n=136)

Factor	Relative rate	Lower 95% CL	Upper 95% CL	Wald statistic	p value
Age at decompensation	1.02	1.00	1.04	3.62 (1)	0.057
Type of decompensation				18.78 (3)	0.0003
Variceal bleeding	1				
Hepatic encephalopathy/ hepatic encephalopathy + other	5.72	2.08	15.72		
Ascites	0.84	0.42	1.69		
Other/Unknown	1.02	0.40	2.60		

The adequacy of the Cox model was checked with the use of partial and martingale residuals (**Figures 4.1.1.15 and 4.1.1.16**). In **Figure 4.1.1.15** it can be seen that the partial residuals are distributed fairly evenly around the zero line, as is expected if the assumption of proportional hazards is met.

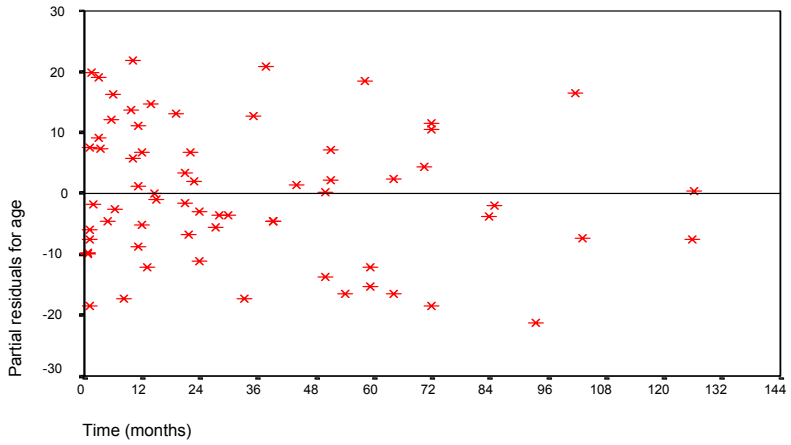


Figure 4.1.1.15. A plot of the partial residuals for age at decompensation from the Cox PH model against survival time for decompensated cirrhosis patients.

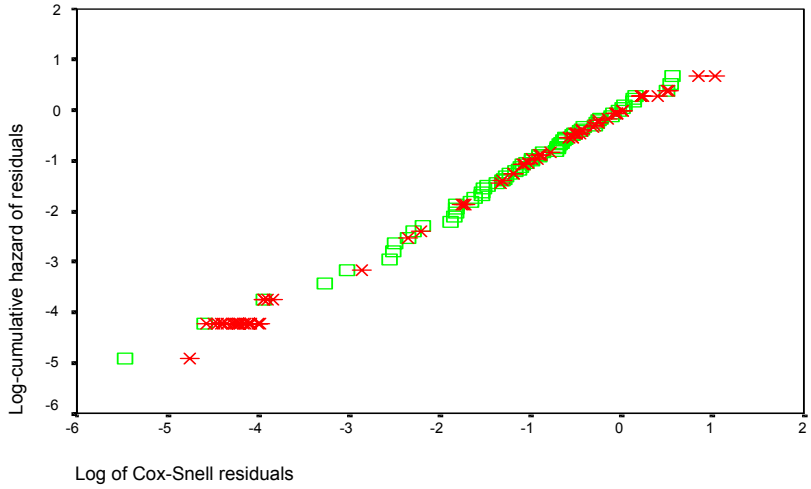


Figure 4.1.1.16 A log-cumulative hazard plot of the Cox-Snell residuals for the Cox PH model fitted for survival time of patients with decompensated cirrhosis.

From the log-cumulative hazard plot of the Cox-Snell residuals (**Figure 4.1.1.16)** it may be inferred that the Cox PH model fits the data well, as the points lie close to a straight line with unit gradient. To check the assumption that time-fixed coefficients are indeed appropriate, the data set was split into 3 data sets , with follow-up censored after 3, 5 and 7 years and models refitted (as described in **Methods 2.1.2.2.2.**). No noticeable trends were found in the estimated coefficients.

Figure 4.1.1.17 depicts sex by survival time for decompensated cirrhotics, with censored values distinguished from uncensored, to test for the possibility of informative censoring (given the 27 losses-to-follow-up). No pattern is evident, thus allowing the inference that it is unlikely that informative censoring is present (see

Methods 2.1.2.2.3.). When a logistic regression model is fitted using backwards stepwise selection (with entry criterion $p<0.05$ and removal criterion $p>0.01$) with age, sex and type of decompensation initially included as explanatory variables and event status as the response status (censored/uncensored), the final model is the null model, again providing evidence that it is satisfactory to use PH regression techniques here.

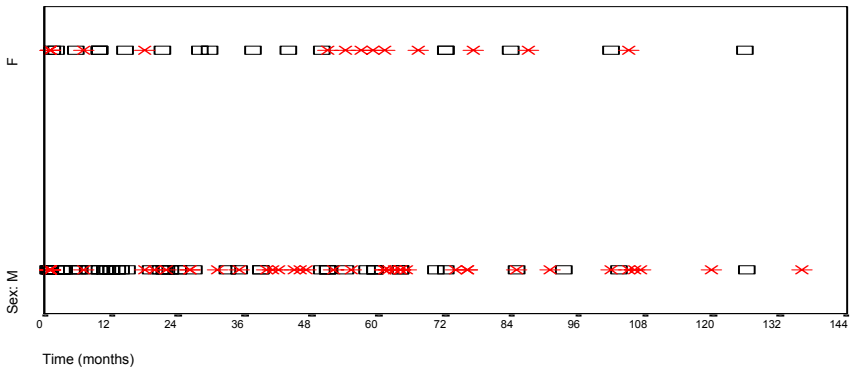


Figure 4.1.1.17 A scatterplot of sex by survival time for decompensated cirrhotics, with censored values distinguished from uncensored.

C. HCC incidence rates

The crude HCC incidence rate in the cirrhosis patients (410 diagnosed as having cirrhosis at the clinic, 12 excluded) was estimated to be 2.3 per 100 person-years. For the 305 compensated cirrhotics, the incidence rate was 2.5 per 100 person-years whereas for the 93 patients with decompensated cirrhosis, it was found to be 1.5 per 100 person-years (this apparently unexpected finding is discussed in **Section 5**).

The overall median time to HCC incidence in cirrhotic patients was found to be 10 years 1 month, with 95% C.I. 9 years 8 months to 10 years 7 months. Overall, 10% of the patients developed HCC at some point during the follow-up period (as detailed in **Section 3.1.**). Sixteen percent of the type B cirrhosis patients developed HCC during follow-up, 14% of those with a combination of a virus and alcohol (although there were only 22 patients in this group), 9% of those with cryptogenic aetiology , 8% of the type C patients and 8% of those with alcohol abuse as aetiology. The cumulative probabilities of the cirrhotic patients remaining tumour-free after 3 years, 5 years and 7 years are provided in **Table 4.1.1.14**, by compensation status. It can be seen that 3 years after diagnosis, 92% of compensated cirrhotics are tumour-free whilst after 5 years and 7 years the corresponding percentages decrease to 87% and 85%

respectively. For those diagnosed with decompensated cirrhosis, the 3-, 5- and 7-year tumour-free rates are 95%, 91% and 91% respectively. From **Table 4.1.1.14** it can been seen that the percentages vary somewhat according to aetiology; at 3 years, only 82% of those with positive HBsAg remain tumour free as compared with between 94% and 96% for all other aetiologies, although the confidence intervals overlap.

Table 4.1.1.14. Estimated cumulative percentages of 398 cirrhosis patients who remain HCC tumour free 3, 5 and 7 years after diagnosis, obtained using Kaplan-Meier analysis

		No. of patients	Cumulative percentage of tumour free patients								
			3 y			5 y			7 y		
			%	Lower 95% CL	Upper 95% CL	%	Lower 95% CL	Upper 95% CL	%	Lower 95% CL	Upper 95% CL
Cirrhotics	compensated	305	92	89	95	87	83	92	85	80	90
	decompensated*	93	95	89	100	91	81	100			
Aetiology	Alcohol	103	94	88	99	92	86	98	72	46	98
	Alcohol+virus*	22	95	85	100	80	59	100			
	HBsAg +	55	82	70	94	76	61	92	76	61	92
	anti-HCV +	162	94	90	98	92	87	96	87	80	95
	No aetiology given/ cryptogenic	56	96	90	100	82	66	98	82	66	98
Sex	male	240	89	95	94	83	77	89	83	77	89
	female	158	97	94	100	95	90	99	90	83	97

* There were no events (i.e. no new tumours) 5 years or more after diagnosis, this may be due to the small numbers in the study by this time

A Cox PH model was fitted using backwards selection with initial variables sex, age, decompensation status and cirrhosis aetiology. The final model contained only age at presentation (an increase of 1 year in age increasing the risk of HCC by about 4%) and sex (with the risk for females being about 1/4 that for males) as significant prognostic factors (**Table 4.1.1.15**). As the rate of censoring was very high (about 90%), however, these estimates should be regarded only as indicative.

Table 4.1.1.15 Estimated relative carcinogenesis rates for factors found to be significantly associated with liver carcinogenesis in a Cox PH model, initially considering sex, age, cirrhosis aetiology and decompensation status at diagnosis (n=398)

Factor	Relative rate	Lower 95% CL	Upper 95% CL	Wald statistic	p value
Sex Female	0.25	0.11	0.58	10.27	0.0013
Age at diagnosis	1.04	1.00	1.07	5.16	0.023

The overall cumulative HCC incidence rates based on the Cox PH model and incidence rates stratified by aetiological group are displayed in **Figures 4.1.1.18** and

4.1.1.19 respectively. The overall estimated cumulative 1-, 3- and 5-year HCC incidence rates (obtained using estimates of the cumulative hazard) were 3% (se 0.9%), 8% (se 1.5%) and 13% (2%) respectively. The estimated cumulative 1-, 3- and 5-year HCC incidence rates in our cirrhosis type B patients were 10% (se 4%), 20% (se 6%) and 27% (se 8%) respectively whilst in our type C patients the corresponding percentages were only 2% (se 1%), 7% (se 2%)and 9% (se 2%) respectively.

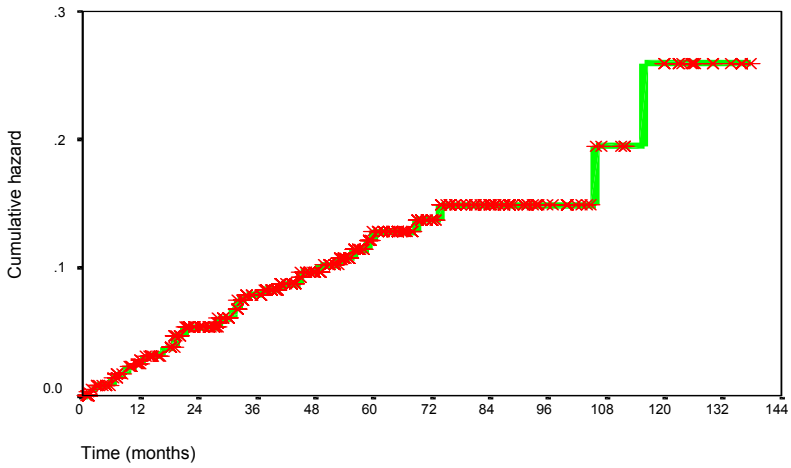


Figure 4.1.1.18. The cumulative incidence rate (cumulative hazard) of hepatocellular carcinoma for 398 cirrhotic patients.

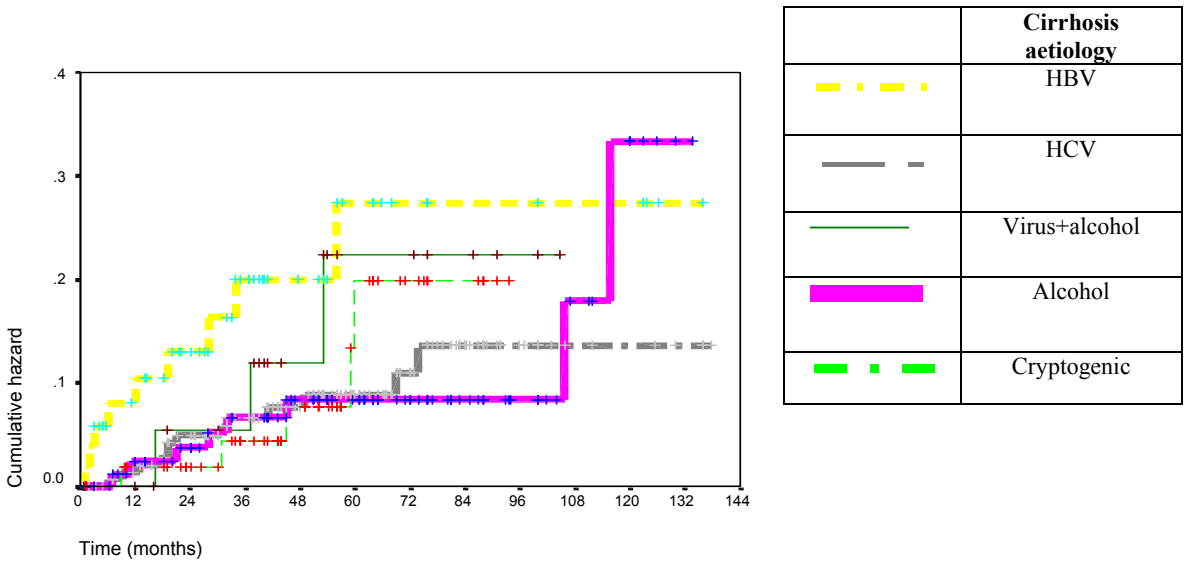


Figure 4.1.1.19. The cumulative incidence rate (cumulative hazard) of hepatocellular carcinoma for 398 cirrhotic patients by aetiological group.

4.1.2. Treatment of type C cirrhosis

For the compensated cirrhosis patients with HCV as aetiology, the cirrhosis was known to be either active (i.e. high aminotransferases) or inactive (i.e. normal levels of aminotransferases) at diagnosis. The former group received *Plaquenil* (hydroxychloroquine) treatment whereas the latter group were not treated and served as controls. There were 162 compensated type C patients in total with aetiology of either solely HCV (151 patients) or a combination of HCV and alcoholism (11 patients). Of these 162 patients, the 52 patients with active cirrhosis were treated for a six-month period whereas the 110 patients with inactive cirrhosis (including 10 of the patients with dual aetiology) were not provided with treatment. A retrospective study of this sub-group of the cirrhosis cohort was undertaken to assess possible effects of treatment with *Plaquenil* on a) time to decompensation b) survival time.

There was no evidence of heterogeneity between groups with regard to their prognostic factors (i.e. age, sex) at diagnosis. Sixty of the patients decompensated at some later stage. Nine of the untreated patients died without evidence of decompensation whilst 4 of the treated patients died without decompensation. The 13 cases of deaths without prior decompensation were treated as censored, as the deaths were presumed to be unconnected to the underlying liver disease. There were 21 deaths due to liver disease in total (13%) over the study period. Of the 52 treated patients, 22 later decompensated (42%) and 4 died (8%) whereas from the 110 untreated patients 38 decompensated (35%) and 17 later died (15%). Fifteen patients developed HCC over the study period (11 non-treated, 4 treated).

Kaplan-Meier PL curves for the estimated time to decompensation by treatment status for HCV cirrhotics are presented in **Figure 4.1.2.1** below. There were 6 missing dates, with corresponding patients omitted from the analysis. There is clearly no evidence of a difference between the two groups at the 5% significance level (log rank test p-value 0.7).

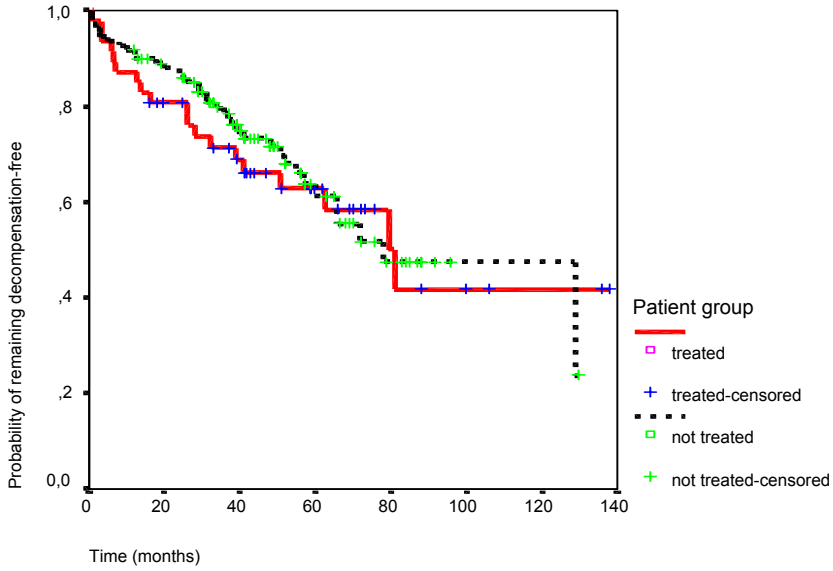


Figure 4.1.2.1 Kaplan-Meier curves for the estimated time to decompensation for 162 type C cirrhotics by treatment status (52 patients with active cirrhosis treated with *Plaquenil*, 110 untreated inactive cirrhotics).

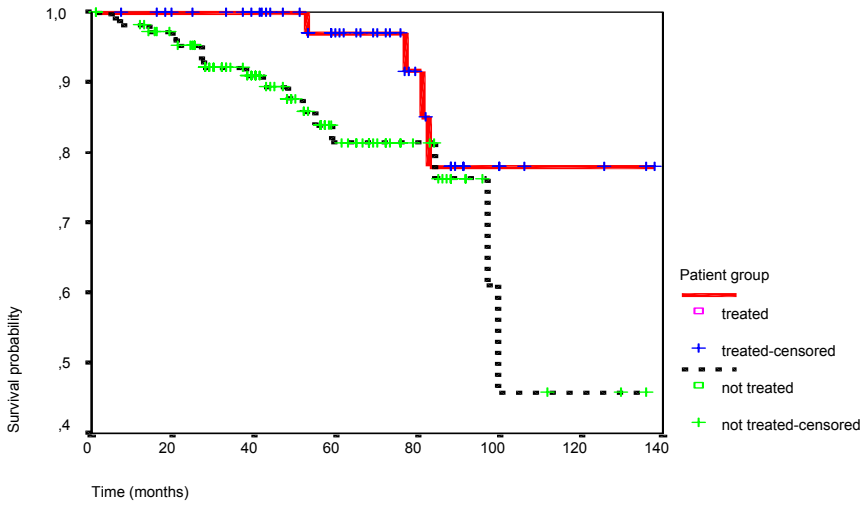


Figure 4.1.2.2. Kaplan-Meier survival curves for 162 type C cirrhotics by treatment status (52 patients with active cirrhosis treated with *Plaquenil*, 110 untreated inactive cirrhotics).

Kaplan-Meier PL estimates of the survival functions by treatment status are presented in Figure 4.1.2.2. above. There is no evidence of a difference in the survival functions between the two groups at the 5% significance level (log rank test p-value 0.06), although from **Figure 4.1.2.2.** it appears that an assumption of proportional hazards may not be appropriate for this group. Application of the generalized Wilcoxon test (which does not assume the hazards are proportional, **Section 2.1.1.4**) provided only weak evidence of a difference in the survival times between the plaquenil-treated and untreated type C patients (test statistic 4.74 on 1 df, p=0.029). By inspection of the graph, it seems that the difference in survival times becomes evident only after long time periods (greater than 8 or 9 years). As no differences in either decompensation or survival rates were detected between the treated and untreated patients, they were regarded as a single group for the present study of cirrhotics (and hence led to the results presented in **Section 4.1.1**.).

It should be noted that treating only deaths after decompensation as being due to liver disease may not be the correct approach if, for example, the treatment has such severe side effects that it may lead to death from causes not involving decompensation. Including all deaths as end-points (not only deaths following decompensation), however, still provides no evidence of a treatment effect. It is likely, however, that there are too few patients available in this analysis to detect a true effect. Using Schoenfeld's formula (**Section 2.1.3**), to detect an effect of $\theta=2$ (i.e. double the relative risk, RR, for the untreated group), with a power of 0.8 and $\alpha=0.05$, at least 356 patients would be needed. This required sample size increases as the RR decreases. Alternatively stated, with the given sample size and RR=0.36, the power of the study is only 59%.

Another point to consider is that the untreated patients were those with inactive cirrhosis whereas the active cirrhotics formed the treatment group. Inactive type C cirrhotics have a better prognosis, on average, than active type C cirrhotics so although the two groups display heterogeneity, any bias is likely to be in favour of the untreated (inactive) cirrhosis patients. A separate point to note is that the heterogeneity between the two groups with regard to the presence/absence of dual cirrhosis aetiology may also play a role in the predictions.

4.1.3. UDCA treatment for PBC patients

The baseline characteristics of the PBC subjects and the univariate analyses involving estimation of the relative risk (RR) of death by prognostic variable are presented in **Table 4.1.3.1**. The vast majority of the PBC patients were female (90%) and middle-aged (mean age 59 years, se 1.1, range 32 to 85 years). Forty three percent of the patients had advanced disease at diagnosis, based on Ludwig staging. The numbers of patients positive for HBsAg and anti-HCV were 2 and 4 respectively, although there were 12 patients whose viral marker statuses were not determined. The missing values were believed to be missing at random.

The overall mean survival time was 117 months, with lower and upper 95% confidence limits of 107 months and 127 months. The baseline factors (i.e. measurements at diagnosis) found to significantly increase the hazard of death using univariate analyses (log rank tests) were being of older age ($p<0.01$), having moderate or severe oedema ($p<0.0001$), having severe disease according to Ludwig staging ($p<0.001$), having low albumin concentration ($p<0.05$), having high prothrombin time ($p<0.05$) and having a high Mayo risk score R ($p<0.05$). For example, the risk of death was 10.2 times higher in patients with Ludwig stage 3 or 4 at baseline than for those with stage 1 or 2. Treating R as a continuous variable in a Cox regression analysis assessing the association of the Mayo risk score with survival, it was found that each unit increase in the risk score increased the risk of death by a factor of 2.8 (95% C.I. 1.9 to 4.2).

The actual survival curve of the UDCA-treated patients, estimated using the Kaplan-Meier product-limit method, and the survival expected for untreated patients using the Mayo model predictions are plotted in **Figure 4.1.3.1.**below. These two curves were found to be significantly different over the first seven year period, with a better actual overall survival of the treated patients than for the simulated control group (observed number of deaths=17, expected number of deaths under Mayo model=39, χ^2= 12.81 on 1 df, $p<0.001$, n=104). From the graphical display, it appears that the difference in the observed survival pattern as compared to the expected survival time increases as the time after diagnosis increases. At 1 year after diagnosis, the observed and expected percentages of survivors coincide at 95%, whereas at 7 years the observed survival rate is 82% as compared to an expected rate of 62%.

Table 4.1.3.1. Relative risk of death (RR) of 104 PBC patients by demographic, clinical, histological and biochemical prognostic variables and also by Mayo risk score

		RR[1]	**Lower 95% CL[2] for RR**	**Upper 95% CL for RR**
Demographic variables				
Median age in years[3] (mean , s.e)	61 (59, 1.1)	3.6	1.2	11.0
Sex, no. of males (%)	10 (10%)	0.4	0.1	1.8
Clinical variable: oedema	**N (%)**			
Minimal oedema	93 (89%)	1		
Moderate oedema	5 (5%)	12.1	3.2	46.4
Severe oedema	6 (6%)	10.0	3.1	32.6
Histologic variable: Ludwig stage[4]	**N (%)**			
Stage 1 or stage 2	59 (57%)	1		
Stage 3 or stage 4	44 (43%)	10.2	2.3	44.8
Biochemical variables[3]	**Median (mean, SE)**			
Total serum bilirubin (mg/dL)	0.90 (1.18, 0.14)	1.7	0.6	4.8
Serum albumin (gm/dL)	4.1 (4.1, 0.06)	0.4	0.1	1.0
Prothrombin time (seconds)	13.0 (13.1, 0.13)	3.6	1.0	12.8
Risk score[3]	**Median (mean, SE)**			
Mayo model risk score	4.64 (4.83, 0.12)	3.2	1.0	9.9

[1] RR=relative risk; Univariate Cox models were used to assess to estimate the RR. The risk for patients in each category is compared to those in the base category (RR=1).
[2]CL=confidence limit
[3]The median was used as a cut-off in dichotomizing continuous variables. Here, the RR given is for the group of patients with values above the median, compared to patients with values below the median (who have RR=1).
[4]One missing value.

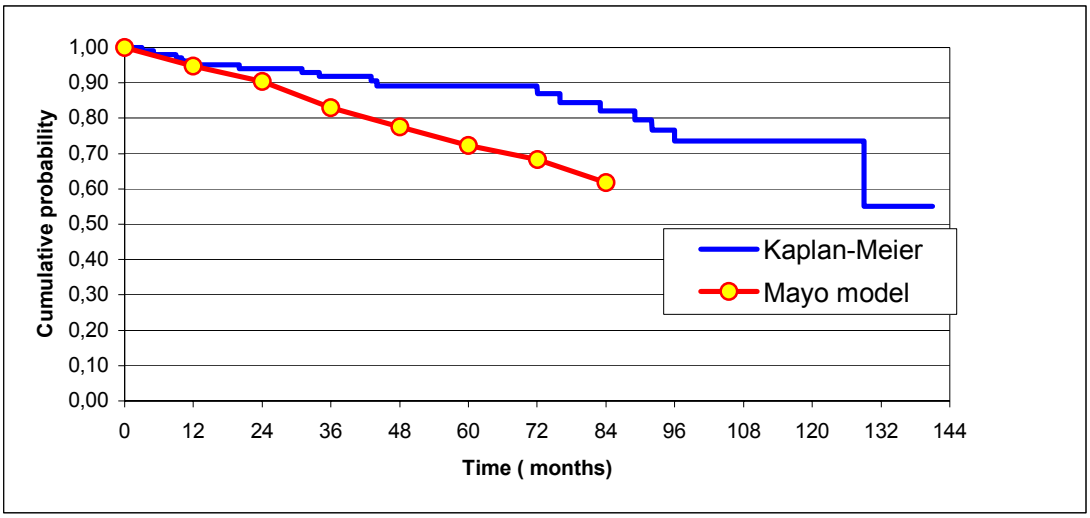


Figure 4.1.3.1. Actual (Kaplan-Meier) survival of 104 UDCA-treated PBC patients and estimated survival as predicted by the Mayo natural history model ($p<0.001$).

In considering the 19 AIC patients as a separate group from the remaining 85 PBC patients, the mean survival time was estimated to be 115 months in each group, with 95% CIs of 97 to 132 months and 104 to 127 months in AIC and non-AIC patients respectively. The mean risk score in the AIC group was 5.03 (with 95% C.I. 4.51 to 5.55) whilst in the non-AIC group it was 4.79 (with 95% C.I. 4.52 to 5.05). **Figure 4.1.3.2** depicts the Kaplan-Meier PL estimate of the survival curve for the AIC patients and the Mayo model curve representing survival of untreated patients. There was found to be significantly higher overall survival in the treated AIC patients than in the simulated control group (observed number of deaths=3, expected number of deaths under Mayo model=11, χ^2= 5.61 on 1 df, $p<0.05$, n=19). As with the overall PBC group, it appears that the difference in the observed survival pattern as compared to the expected survival time increases as the time after diagnosis increases. At 1 year after diagnosis, the observed percentage of survivors is 95%, compared to a prediction of 94% whereas at 7 years the observed survival rate is 80% as compared to an expected rate of 58%. In **Figure 4.1.3.3,** which depicts the Kaplan-Meier PL estimate of the survival curve for the non-AIC patients and the Mayo model curve representing the survival pattern of a similar group of untreated patients, the same general pattern is seen as in **Figures 4.1.3.1 and 4.1.3.2**. At 1 year after diagnosis, the observed percentage of survivors is 95% as predicted, whereas at 7 years the observed survival rate is 83% as compared to an expected rate of 63%. There was found to be significantly higher overall survival in the treated non-AIC patients than in the

simulated control group (observed number of deaths=14, expected number of deaths under Mayo model=29, χ^2= 7.54 on 1 df, p<0.01, n=85).

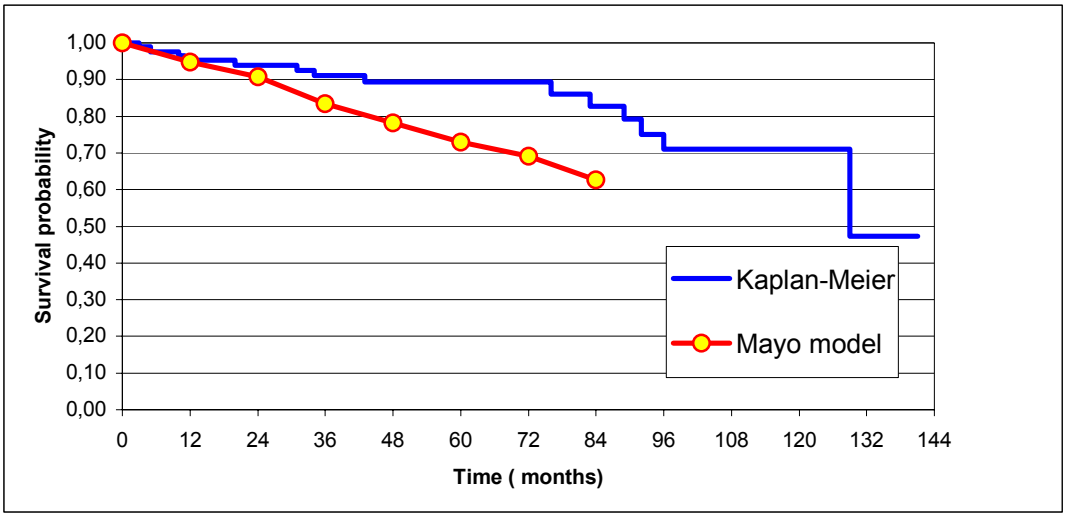


Figure 4.1.3.2. Actual (Kaplan-Meier) survival of 85 non-AIC UDCA-treated PBC patients and estimated survival as predicted by the Mayo natural history model (p<0.01).

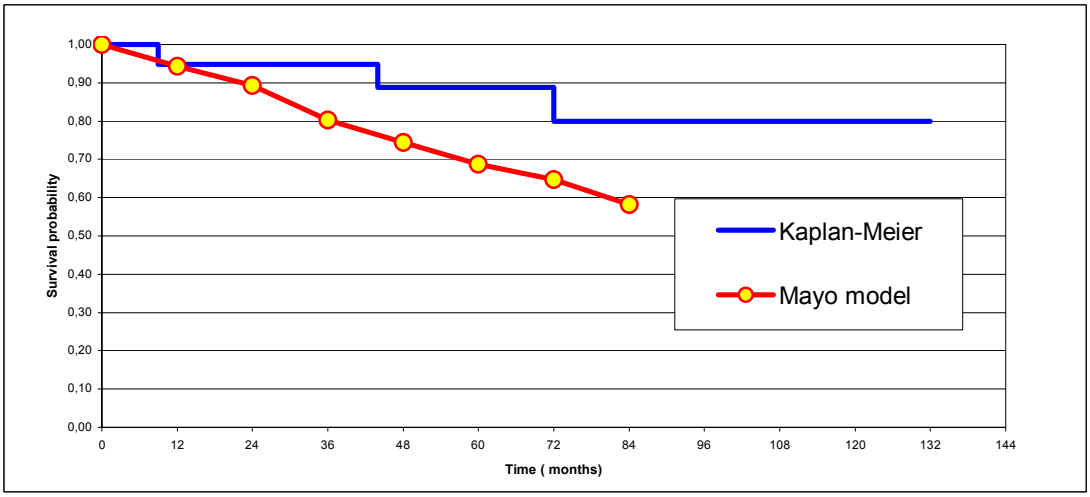


Figure 4.1.3.3. Actual (Kaplan-Meier) survival of 19 UDCA-treated PBC patients with AIC and estimated survival as predicted by the Mayo natural history model (p<0.05).

The Mayo model risk score was very highly correlated with bilirubin, albumin, prothrombin time and age at diagnosis (p<0.0001 in all cases using Spearman rank correlation coefficients), as one might expect given that the combination of these variables led to the creation of the risk score. Multivariate Cox regression analysis was therefore undertaken with the following three variables initially considered: Ludwig stage, AIC status and risk score. The Cox model resulting from the stepwise

selection procedures contained the risk score (RR 2.5, 95%C.I. 1.6 to 3.9, p<0.0001) and Ludwig stage (III,IV versus I,II, RR 5.7, 95% C.I. 1.2 to 26.1,p=0.025). The stability of the model was investigated using bootstrap analysis (with B=2000 replicates). In **Table 4.1.3.2** it can be seen that the estimated regression coefficients and their SEs prior to resampling differ from the bootstrap results, with estimated biases 0.85 and 0.12 for Ludwig stage and R respectively. The medians and the bootstrap mean of the regression coefficent replicates were not in close agreement for either variable. Histograms of the empirical distribution of replicated regression coefficients for each variable with a smoothed density estimate are provided in **Figure 4.1.3.4,** from which it can be seen that the distribution appears distinctly non-normal for the Ludwig stage variable, with the distribution of the coefficients for R appearing somewhat positively skewed. Using jackknife after bootstrap techniques to assess the influence of each of the observations on the bias, there were found to be six influential points, two of these being highly influential (with absolute relative influence on the bias of 6.9 and 6.6 for the Ludwig variable). When the Cox model was refitted omitting the six influential observations, although the RR were not much changed the confidence intervals for the Ludwig stage variable became extremely wide, indicating instability of the coefficient. The same occurred when the model was refitted leaving out only the two most influential observations.

When the Cox procedure was repeated considering the five variables used in the creation of the Mayo model, the final survival model for our PBC group contained only the age at diagnosis (RR 1.1, 95%C.I. 1.0 to 1.1, p=0.038) and $\log_e$(albumin) concentration (RR 0.002, 95% C.I. 0.000 to 0.046, p<0.0001). Bilirubin concentration, oedema status and prothrombin time were not found to be significant predictors of survival when considered jointly with the other Mayo model variables. Again, bootstrapping techniques indicated the instability of the model coefficients.

Table 4.1.3.2. Bootstrap estimates of regression coefficients and standard errors based on the Cox survival model for PBC patients with R and Ludwig stage as prognostic variables (2000 replicates)

Variable	Regression Coefficient		Bootstrap regression coefficient		Estimated bias	Median & 90% empirical confidence limits			Median & 90% BCa confidence limits		
	$\widehat{\beta}$	$se(\widehat{\beta})$	$\widehat{\beta}$*	$se(\widehat{\beta}$*)		5%	50%	95%	5%	50%	95%
Ludwig stage:III/IV	1.742	0.777	2.564	2.561	0.822	0.527	1.753	9.055	0.364	1.735	8.835
Risk score, R	0.921	0.224	1.035	0.386	0.114	0.603	0.983	1.629	0.487	0.862	1.405

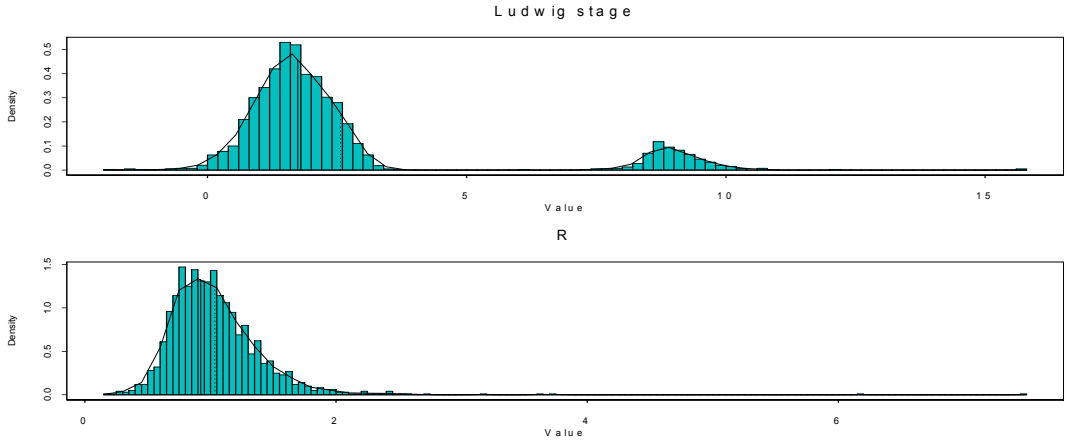


Figure 4.1.3.4. Histograms of the empirical distributions of parameter replicates for the Cox survival model for PBC patients (2000 replicates); the model contained the variables Ludwig stage (III or IV versus I or II) and R, the Mayo risk score.

4.1.4. Natural history of HCC

Log rank test results are presented in **Table 4.1.4.1**. There was significant evidence of a difference in the survival distributions at the different levels of the following variables: *tumour* (p=0.033), *Okuda index* (p=0.0002), *AgHBe* (p=0.018) and *albumin* (p=0.030) and marginal evidence of a difference in the survival functions for the variable *prothrombin time* (p=0.062).

Table 4.1.4.1. Estimated median survival times of 48 HCC patients by demographic, clinical and biochemical prognostic factors

Factor	**Groups**	**Median survival time in months (95% C.I.)**	**Log rank statistic (d.f.)**	**p-value**
Sex	Male	6 (5,7)	1.78 (1)	NS
	Female	3 (2,4)		
Age (years)	<68	6 (4,8)	1.18 (2)	NS
	68-72	5 (1,9)		
	>72	6 (1,11)		
Place of residence	Iraklion	5 (2,8)	1.13 (4)	NS
	Rethymnon	5 (2,8)		
	Lassithi	8 (3,13)		
	Hania	6 (3,9)		
	Other	3 (-,-)		
Tumour	One	6 (4,8)	4.60 (1)	0.032
	> one	3 (1,5)		
Okuda stage	I	16 (-,-)	17.01 (2)	0.0002
	II	7 (4,10)		
	III	2 (1,3)		
Cirrhosis	Present	6 (4,8)	0.12 (1)	NS
	Absent	7 (4,10)		
Ascites	Present	6 (4,8)	1.65 (1)	NS
	Absent	6 (0,12)		
Hepatitis indicators:				
HCV	Positive	7 (4,10)	0.21 (1)	NS
	Negative	6 (5,7)		
HBsAg	Positive	4 (1,7)	2.01 (1)	NS
	Negative	6 (5,7)		
AgHBe	Positive	1 (-,-)	5.88 (1)	0.018
	Negative	6 (4,8)		
anti-HBc	Positive	5 (3,7)	4.18 (1)	0.041
	Negative	8 (4,12)		
anti-HBs	Positive	6 (4,8)	0.23 (1)	NS
	Negative	6 (4,8)		
AFP (ng/ml)	<101	6 (5,7)	2.95 (2)	NS
	101-699	5 (2,8)		
	>699	7 (2,12)		
Bilirubin (mg/dL)	<1.5	6 (4,8)	4.41 (2)	NS
	1.5-3.0	6 (5,7)		
	>3.0	3 (1,5)		
Prothrombin time (s)	<13.5	6 (3,9)	5.58 (2)	NS
	13.5-14.49	5 (2,8)		
	>14.49	2 (1,3)		
Albumin (g/L)	<31	2 (0,4)	7.04 (2)	0.030
	31-36	8 (4,12)		
	>36	6 (2,10)		

The Kaplan-Meier estimated survival curves for the prognostic factors are given in **Figures 4.1.4.1** to **4.1.4.6** below. As the Kaplan-Meier curves for prothrombin time indicated the possibility of non-proportional hazards, the generalized Wilcoxon test was applied (**Methods 2.1.1.4.**). The generalized Wilcoxon statistic for the *prothrombin time* factor was 8.37 (p=0.015). For the other factors, there was no indication of non-proportional hazards so the log-rank test was assumed appropriate.

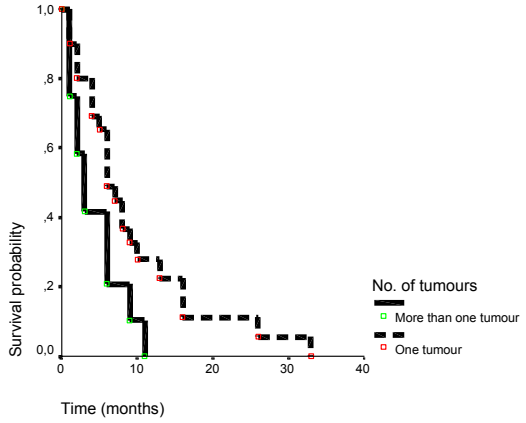


Figure 4.1.4.1. Kaplan-Meier estimate of the survival function from the time of diagnosis, for 48 untreated HCC patients, by the number of tumours present.

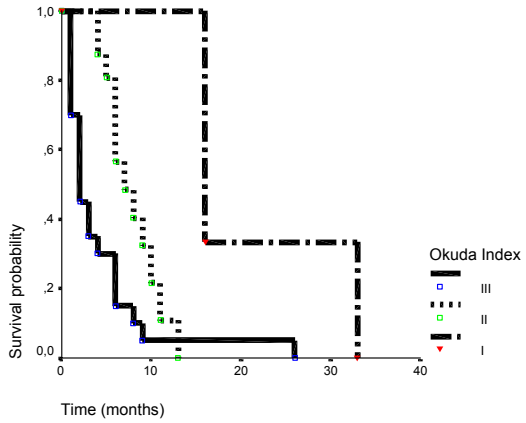


Figure 4.1.4.2. Kaplan-Meier estimate of the survival function from the time of diagnosis, for 48 untreated HCC patients, by Okuda staging.

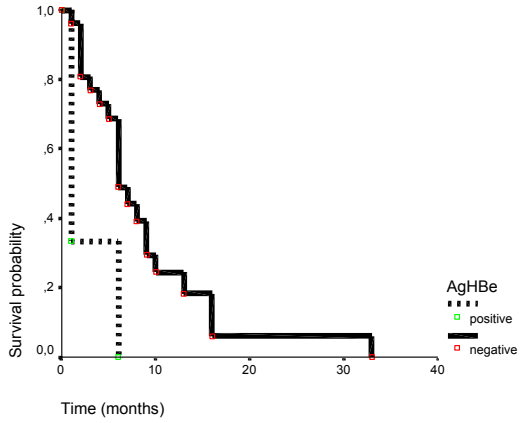


Figure 4.1.4.3. Kaplan-Meier estimate of the survival function from the time of diagnosis, for 48 untreated HCC patients,, by hepatitis B e antigen (AgHBe) status.

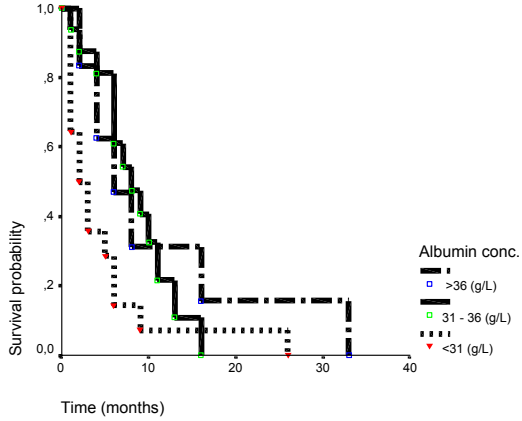


Figure 4.1.4.4. Kaplan-Meier estimate of the survival function from the time of diagnosis, for 48 untreated HCC patients, by albumin concentration.

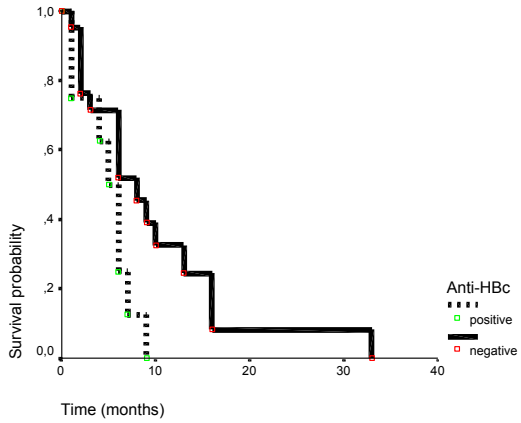


Figure 4.1.4.5. Kaplan-Meier estimate of the survival function from the time of diagnosis, for 48 untreated HCC patients, by antibody to hepatitis core antigen (anti-HBc) status.

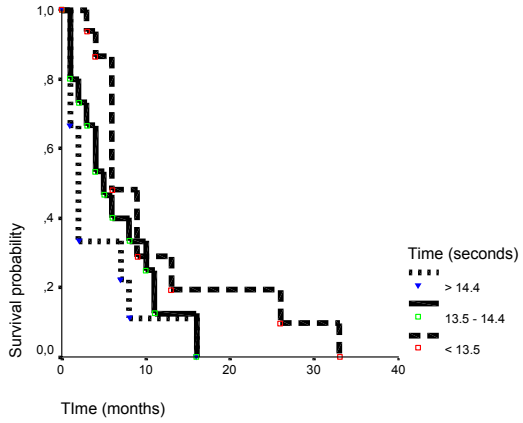


Figure 4.1.4.6. Kaplan-Meier estimate of the survival function from the time of diagnosis, for 48 untreated HCC patients, by prothrombin time (seconds).

Results from the individual Cox's proportional hazards models fitted using the forced-entry method are presented in **Table 4.1.4.2**. The contrasts used are indicator contrasts, comparing each level of the factor to the lowest/negative level.

Table 4.1.4.2. Univariate Cox model relative risks (RR) in the untreated HCC patients for the following significant prognostic factors: number of tumours, Okuda index, HBeAg status, albumin concentration, prothrombin time

Variable	χ^2* (d.f.)	p-value	R.R.	95% C.I.
>1 tumour	4.04 (1)	0.044	2.12	(1.00 , 4.48)
Okuda index	14.83 (2)	0.0006		
II			4.17	(0.88 , 19.85)
III			9.72	(2.17, 43.46)
Hepatitis indicators:				
HBeAg	4.78 (1)	0.029	3.78	(1.05 , 13.60)
Concentrations:				
Albumin	6.17 (1)	0.013	0.89	(0.81 , 0.98)
Prothrombin time (s)	5.54 (1)	0.019	1.54	(1.07 , 2.20)

*Score test statistic

The R.R. of dying for those with *Okuda* index II is 4.2 times that for those with *Okuda* index I whilst the R.R. of dying for those with *Okuda* index III is 9.7 times that for those with *Okuda* index I. The R.R. for those with *AgHBe* positive is approximately 3.8 times that for those with *AgHBe* negative. For the *albumin* variable, a unit increase in albumin concentration results in an 11% decrease in the hazard rate.

As the number of patients was relatively small and, in addition, the data set contained some missing values, a multivariable model was not fitted. Therefore, only limited inferences can be drawn. It appears that:

- having multiple tumours is associated with a higher risk of death as compared to having one tumour
- those with Okuda indices of II and III have a higher risk of death than those classified as Okuda I.
- the hepatitis indicators AgHBe and anti-HBc being positive may be an indication that the risk of death is higher.
- those patients with higher albumin concentrations have a lower instantaneous risk of death
- as the prothrombin time ('quick' time) increases, the hazard increases**.**

4.1.5. Treatment for HCC patients

4.1.5.1. Octreotide

Table 4.1.5.1.1. shows the estimated average survival times and the percentages of patients surviving 6- and 12-months by each prognostic variable. Thirty seven percent of non-treated and 75% of those treated were alive at six months. At 12 months, the corresponding percentages were 13% and 56% and at 24 months 3% and 20%.

Table 4.1.5.1.1. Estimated median survival times and cumulative survival percentages at 6- and 12-months by prognostic factor for 58 HCC patients: 28 treated with octreotide and 30 untreated controls

Factor	**No. of patients**	**Median survival (months)**	**Percentage surviving at 6 months**	**Percentage surviving at 12 months**	**Log rank p-val.**
Treatment					0.002
Octreotide	28	13.0	75	56	
None (controls)	30	4.0	37	13	
Cirrhosis					0.029
Present	47	6.0	47	31	
Absent	11	8.0	91	40	
Okuda stage					0.020
I	5	16.0	100	100	
II	23	7.0	55	40	
III	30	7.0	56	26	
Sex					NS
Male	48	7.0	58	34	
Female	10	4.0	40	30	
Tumour size					NS
Small	4	11.0	75	50	
Medium	11	9.0	89	44	
Large	21	4.0	44	31	
Multiple	22	7.0	56	39	
Age (years)					NS
<67	19	7.0	53	32	
67-72	20	5.0	50	22	
>72	19	8.0	63	47	
Place of residence					NS
Heraklion	21	7.0	62	38	
Rethymnon	23	7.0	56	37	
Lassithi	8	5.0	50	25	
Hania	4	1.0	25	25	
Other/unknown	2	6.0	50	0	
AFP **(ng per ml)**					NS
<90	19	8.0	58	31	
90-620	20	6.0	45	25	
>620	19	7.0	63	46	
Bilirubin (mg%)					NS
<1	16	13.0	65	53	
1-6	22	7.0	52	33	
>6	20	7.0	60	27	
Serum albumin (g/L)					0.016
<3	18	3.0	36	14	
3-4	21	8.0	68	29	
>4	19	14.0	69	69	
Viral markers					NS
Anti-HCV	31	8.0	61	39	
HBsAg	14	4.0	36	21	
HbsAg and anti-HCV	2	6.0	50	0	
Negative	11	8.0	64	42	

The log rank test provided evidence of significant differences for the following factors:

a) *treatment* (p=0.0024, df=1). The median survival time for those patients who received treatment was 13 months (se=1.90, with 95% C.I. 9 to 17 months) whilst for those who did not receive treatment it was 4 months (se=1.10 and 95% C.I. 2 to 6 months). The Kaplan-Meier PL estimate of the survival function by treatment status is provided in **Figure 4.1.5.1.1**.

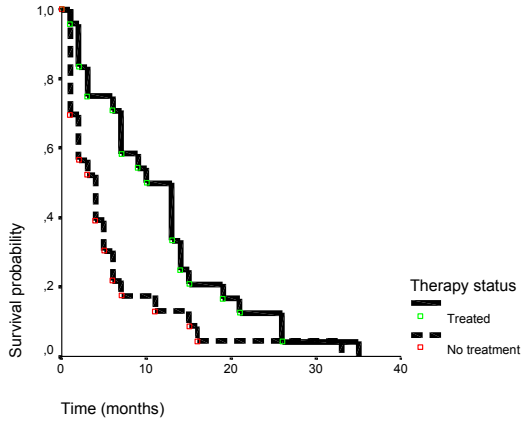


Figure 4.1.5.1.1. Kaplan-Meier estimates of the survival function by treatment status (24 patients treated with octreotide, 23 untreated controls).

The NNT at 12 months is estimated to be 2.3 (i.e. giving patients treatment would lead to 1 extra survivor at 1 year for every 2.3 patients treated) with 95% C.I. 1.5 to 4.9 patients needing to be treated for one extra survivor after one year.

b) *cirrhosis* (p=0.0285, df=1). Those patients with cirrhosis appear to have a higher instantaneous risk of death than thse not suffering from cirrhosis, even though the number of non-cirrhotics in the sample is small (11 subjects). **Figure 4.1.5.1.2.** displays the Kaplan-Meier PL estimate of the survival function by treatment status only for the cirrhotics (n=47, p=0.0114, df=1). As 81% of the HCC patients were also cirrhotics, it is not surprising that the curves closely resemble those of **Figure 4.1.5.1.1**.

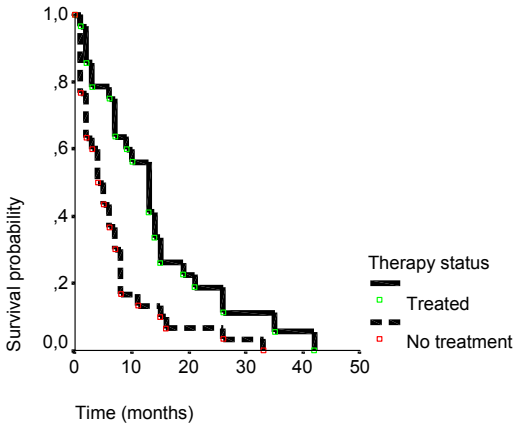


Figure 4.1.5.1.2. Kaplan-Meier PL estimates of the survival function by treatment status for cirrhotic patients with HCC (24 patients treated with octreotide, 23 untreated controls).

c) *albumin* (p=0.0161, df=2). Higher albumin concentrations were associated with higher survival rates

d) *Okuda* stage (p=0.020,df=2). Okuda I patients have a median survival time of 16 months (SE 9.0 months) whereas Okuda II and Okuda III patients have median survival times of 7 months (SE 2.2 and 0.8 respectively). Further logrank tests were performed to test the effectiveness of the drug whilst controlling for Okuda I/II versus III patients and also small/medium versus large/multiple tumours. The subgroup results are presented in **Table 4.1.5.1.2.**, from which it can be seen that treatment remains effective even after controlling for Okuda staging and tumour size.

Table 4.1.5.1.2. Comparison of survival distributions of treated versus untreated patients, controlling for Okuda grouping (log-rank p = 0.013) and tumour size (log-rank p = 0.009)

	Median survival times in months (SE)	
	Treated	Untreated
Okuda I or II	13 (1.93)	6 (1.49)
Okuda III	9 (2.51)	3 (1.73)
Small/medium tumour	19 (2.83)	11 (4.38)
Large/multiple tumour	13 (2.12)	4 (0.65)

In order to assess possible influences of other variables, Cox PH models were fitted to the data using stepwise regression procedures using likelihood ratio (LR) tests for overall model-fitting with score statistic entry criterion $p<0.05$ and Wald statistic removal criterion $p>0.1$ for each variable at each step. The most suitable model contained the variables *treatment, log (AFP)*, *albumin*, *cirrhosis*, and *tumour size*, as

shown in **Table 4.1.5.1.3.** Treated patients had an instantaneous risk of death 0.38 times that of those who did not undergo treatment, accounting for the other prognostic variables. As the concentrations of albumin and AFP increased, the hazard decreased. A unit increase in the albumin concentration decreased the hazard rate by about 10%, all other covariates remaining unchanged. The relative risk for those with cirrhosis was 5.5 i.e. the estimated relative risk of dying was 5.5 times greater for those with cirrhosis, adjusting for the other covariates. Having medium, large and multiple tumours was associated with a having a higher risk of death compared to having small tumours.

Table 4.1.5.1.3. Multivariate Cox PH regression analysis relative risks (RR) for the significant variables for HCC prognosis

			95% C.I. for $e^{\beta j}$	
Variable	**p-value**	**RR ($e^{\beta j}$)**	**Lower limit**	**Upper limit**
Treatment	0.010	0.38	0.18	0.79
Log AFP	0.070	0.88	0.76	1.01
Cirrhosis	0.002	5.47	1.90	15.78
Albumin	0.012	0.90	0.83	0.98
Tumour size	0.108			
Medium vs small	0.558	1.50	0.39	5.80
Large vs small	0.042	3.78	1.05	13.31
Multiple vs small	0.196	2.28	0.65	7.91

4.1.5.2. Long-acting somatostatin analogues

From **Table 3.3.3.1** (**Section 3.3.3**) it appears that the two patient groups display prognostic homogeneity for all discrete factors. Also, no evidence was found of differing age distributions between the two groups (Mann-Whitney test). The estimated 12-month survival rates for treated patients and historical controls were 61% (with 95% C.I. 43% to 80%) and 30% (with 95% C.I. 10% to 50%) respectively. In **Table 4.1.5.2.1**, product-limit estimates of six- and twelve- month survival percentages and corresponding confidence intervals are presented for the treated patients. The median age of the treated patients (71 years) was taken as the cut-off for the categorization in **Table 4.1.5.2.1**. It should be borne in mind that some estimates provide only a vague indication of the true proportions as they may be based on very small numbers in the sub-groups: see **Table 3.3.3.1** for the actual numbers involved.

Table 4.1.5.2.1. Estimated cumulative 6- and 12-month survival percentages by prognostic factor for 32 long-acting somatostatin-treated HCC patients.

	Percentage surviving six months (95% CI)	**Percentage surviving twelve months (95% CI)**
Overall	**74 (58, 89)**	**61 (43, 80)**
Sex		
Male	71 (53, 89)	56 (36, 77)
Female	83 (54, 100)	83 (54, 100)
Okuda stage*		
I	100	90 (71, 100)
II or III	60 (39, 82)	46 (21, 70)
Child-Pugh*		
A	82 (67, 98)	67 (47, 87)
B or C	50 (15, 85)	50 (15, 85)
Cirrhosis		
present	71 (53, 89)	67 (47, 87)
absent	83 (54, 100)	83 (54, 100)
BCLC		
A_1 or A_2	100	89 (68, 100)
B	63 (21, 100)	42 (0, 85)
C	57 (32, 83)	50 (24, 76)
Viral markers		
HCV+	75 (45, 100)	56 (17, 95)
HBV+	76 (47, 100)	76 (47, 100)
Negative	69 (44, 94)	52 (24, 80)
Therapy label		
Octreotide LAR	72 (49, 95)	64 (39, 90)
Lanreotide	75 (54, 96)	53 (39, 86)
Age		
Less than 71 years	72 (48, 95)	48 (14, 81)
At least 71 years	75 (54, 96)	69 (46, 91)

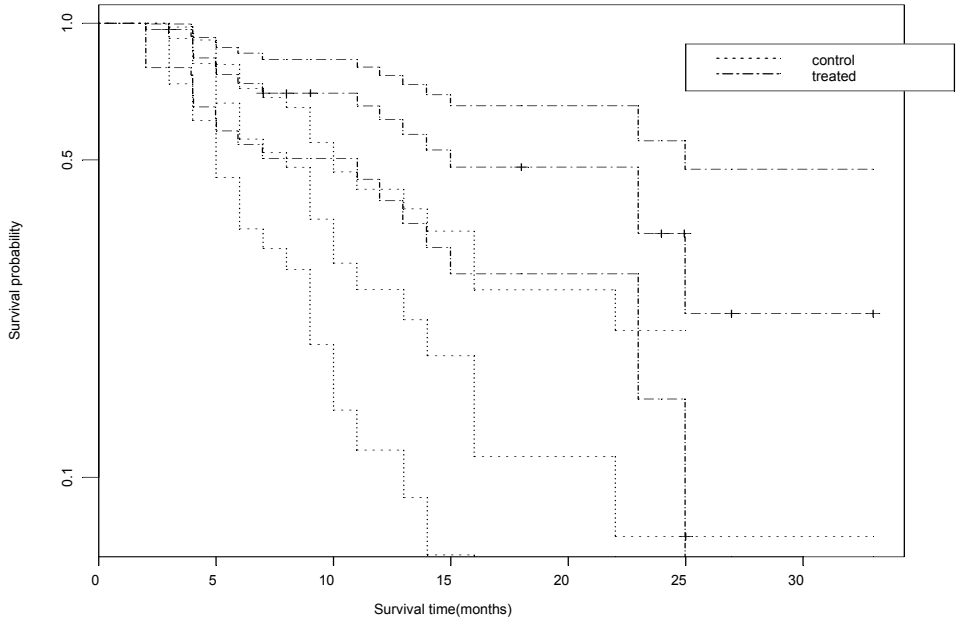


Figure 4.1.5.2.1. Kaplan-Meier PL estimates of the survivor function for 52 HCC patients by treatment status (32 treated with a long-acting somatostatin analogue, 20 untreated controls), with 95% confidence intervals.

From **Figure 4.1.5.2.1.**, it appears that there is a difference in the survival functions between the two groups, with the untreated patients having a higher estimated risk of death at each time point. The corresponding log rank test (1 df, test statistic 4.58) had a p-value of 0.032, leading to rejection of the null hypothesis of no difference in the risk of death for treated versus untreated patients at the 5% significance level. The estimated median survival times and 95% C.I. were 15 months (with 95% C.I. 6 to 24 months) and 6 months (95% C.I. 2 to 10 months) for the treated and untreated patients respectively. The NNT at 12 months is 3.2 (i.e. giving patients treatment would lead to 1 extra survivor at 1 year for every 3.2 patients treated), with 95% C.I. (1.7, 29 patients needing to be treated).

Forwards and backwards Cox regression models were fitted, using likelihood ratio (LR) tests for overall model-fitting with entry criterion $p<0.05$ and removal criterion $p>0.1$ for each variable at each step. The results of the Cox regression analysis indicate that the ratio of the estimated hazard rates for untreated patients compared to treated patients is 3.1 (95% C.I. 1.5 to 6.4), having adjusted for possible differences in Okuda indices i.e. for any particular time interval, an untreated patient is estimated to

be at 3.1 times the risk of death of an treated patient (N.B. a ratio of 1 would imply equal risk). It was found that it was only necessary to adjust for the binary Okuda factor in the final model, with a hazard 5.2 times greater for those being classified as Okuda II (including the two Okuda III treated subjects in this group) as compared to Okuda I patients, with 95% C.I. (2.2, 12.3).

The assumption of proportional hazards was checked using a log-minus-log plot of the estimated survival function (ln[-ln S(t)] vs t). The curves of the two groups of patients appear approximately parallel (**Figure 4.1.5.2.2.)**.

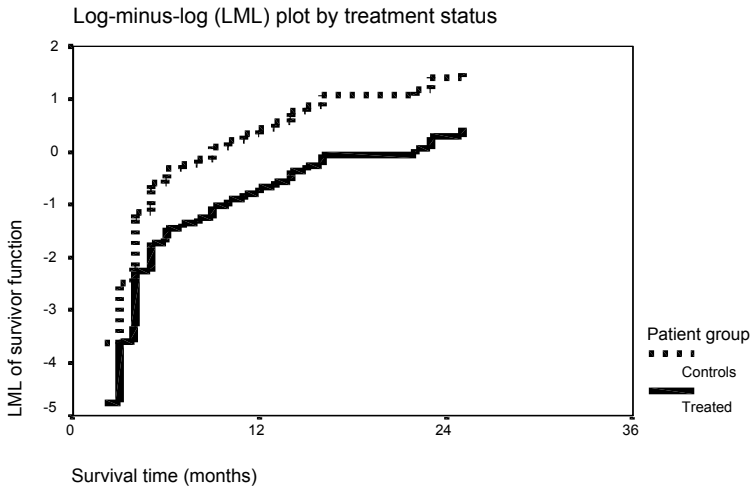


Figure 4.1.5.2.2. Log-minus-log (LML) plot by treatment status (32 HCC patients treated with a long-acting somatostatin analogue, 20 untreated HCC controls).

4.2 Classification of ascitic patient groups using biochemical data

4.2.1 Distinguishing between non-malignant cirrhotic ascites and malignant ascites

In the **group A** data, the ascitic fluid to serum ratios of ceruloplasmin (CER), a_2 – macroglobulin (AMG), haptoglobin (HAP), a_1-antitrypsin (AAT), a_1-acid glycoprotein (AAG), transferrin (TRF) and the immunoglobins IgA, IgG and IgM were found to be highly correlated (r>0.5). **Table 4.2.1.1.** depicts the median and upper (P75) and lower (P25) quartiles of the ratios for which there was found to be a significant difference between patients with malignant neoplasms and those with cirrhosis at the adjusted 5% significance level.

Table 4.2.1.1. Medians, lower and upper quartiles of biochemical parameter ratios in 23 patients with non-malignant cirrhosis and 27 patients with malignant neoplasms where significant differences were detected between the two groups at a univariate level.

Biochemical parameter ratios (ascitic fluid:serum)	**Cirrhosis (n=23)**			**Malignant neoplasm (n=27)**			**Sensitivity (%)**	**Specificity (%)**
	P25	**Median**	**P75**	**P25**	**Median**	**P75**		
Protein	0.21	0.27	0.47	0.60	0.73	0.81	93	87
Albumin	0.12	0.21	0.31	0.62	0.77	0.90	89	87
Lactate dehydrogenase (LDH)	0.36	0.50	0.58	0.76	1.25	2.42	85	83
Ferritin	0.23	0.35	0.60	0.79	1.15	1.98	81	78
Immunoglobulin IgG	0.11	0.30	0.36	0.40	0.69	0.84	81	78
Ceruloplasmin (CER)	0.18	0.36	0.44	0.40	0.55	0.63	74	70
a_2 – macroglobulin (AMG)	0.12	0.15	0.22	0.27	0.32	0.42	81	78
a_1-antitrypsin (AAT)	0.15	0.22	0.38	0.34	0.57	0.70	70	65
a_1-acid glycoprotein (AAG)	0.12	0.27	0.37	0.40	0.59	0.73	78	74
Transferrin (TRF)	0.23	0.34	0.45	0.45	0.63	0.74	78	74
Interleukin 8 (Il-8)	1.53	1.91	2.22	3.92	6.62	10.31	85	83

In **Table 4.2.1.1.** it can be seen that no unique variable displays 100% sensitivity or 100% specificity. There was no evidence of a difference between the two groups with respect to the average ratios of C3 (medians 0.66 and 0.79 for the cirrhosis and

malignant neoplasm groups respectively), C4 (medians 0.56 and 0.72), IL-1a (medians 0.68 and 1.06), IL-2 (medians 0.61 and 0.69), IL-6 (medians 15.52 and 19.21), CRP (medians 0.43 and 0.40), IgA (medians 0.27 and 0.50), IgM (median 0.28 and 0.40), haptoglobulin (median 0.22 and 0.25), IL-1b (medians 0.61 and 0.63), or TNF-a (medians 1.63 and 2.01).

Figure 4.2.1.1. is a schematic representation of the results of the recursive partitioning model obtained when initially entering all variables found to be significant at a univariate level. The most significant parameter was found to be the albumin ascitic fluid:serum ratio, split at a value of 0.392 which separated 20 cirrhotics from the remaining 27 patients with malignant neoplasms and 3 cirrhotics (with deviance 68.99). The next most important parameter was deemed to be the IL-1a ascitic fluid: serum ratio, split at 2.17, which separated 25 of the 27 subjects with malignant neoplasms from the 5 remaining subjects (with deviance 6.73). The final split again involved the albumin ratio (and separated the two subjects with malignant neoplasms from the three cirrhotics, with zero deviance). The model had a misclassification rate (and residual deviance) of 0 i.e. 100% correct classification of subjects into the two disease groups using only the biochemical parameter albumin and IL-1a ascitic fluid: serum ratios.

Applying the rules is simple, and can be illustrated as follows: a patient with an albumin ratio of 0.45 and IL-1a ratio of 2.0 would be predicted as being in the malignant neoplasm group (no other biochemical parameters are required). **Figure 4.2.1.2.** is a scatter plot of the albumin ratios against IL-1a ratios for each disease group, from which it can be seen that the patients can be divided completely into the 2 disease groups using only the 2 biochemical parameter ratios predicted by the recursive partitioning model. As can be seen from the scatter plot, the vast majority of subjects (47 out of 50) can, in fact, be distinguished on the basis of only the albumin ratio.

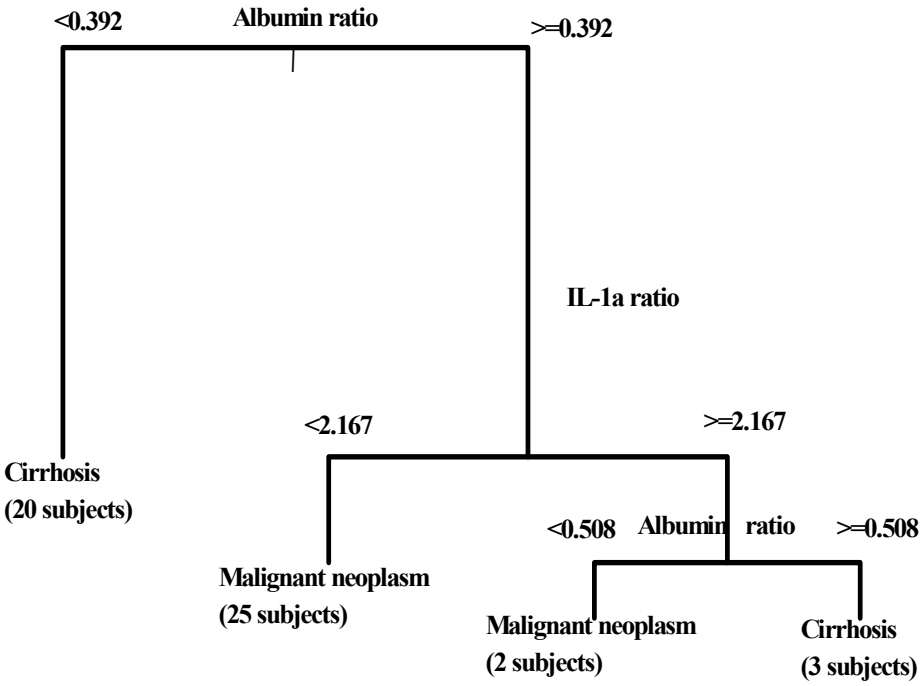


Figure 4.2.1.1. A graphical display of the recursive partitioning rules which discriminate between patients with cirrhosis and those with malignant neoplasms.

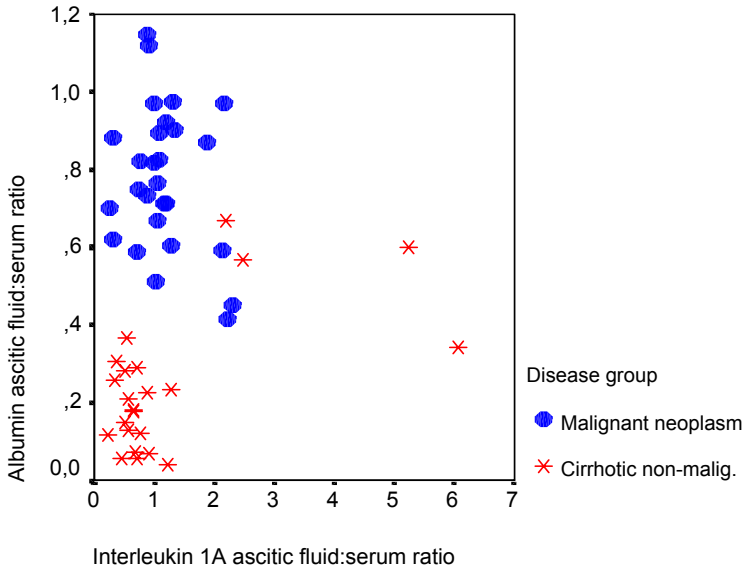


Figure 4.2.1.2. Scatterplot of albumin ratios by interleukin-1a ratios by disease group (27 ascites patients with malignant neoplasms and 23 ascites patients with non-malignant cirrhotic effusions).

A second approach taken was to initially include in the model only those variables found to be significant in the univariate tests and to apply more stringent conditions to the recursive partitioning (i.e. a minimum split of 10 and a minimum node size of 5). Only the albumin ratio was found to be a significant predictor, with an initial split of 0.392, a second split of 0.683, a final split of 0.595 a residual mean deviance of 0.264, and a misclassification error rate of 0.06. This misclassification rate is precisely that expected from the above scatter diagram (1-47/50).

4.2.2 Distinguishing between malignant exudates, non-malignant exudates and transudates

In the measurements taken from **group B** ascites patients (described in **Section 3.4**), many of the serum and ascitic fluid measurements were highly correlated. The seven significant variables in the serum measurements using one-way ANOVA were found to be tumor necrosis factor –alpha (TNF-α), complement factor C3, complement factor C4, interleukin-1a (IL-1a), HAP and the acute phase proteins AAG and AAT. Summaries are presented in **Table 4.2.2.1**. The groups between which the significant differences lie are provided by the SNK contrasts, from which it appears that in the main it is the two exudate groups that differ from the transudate group. For C4 and AAT there also appears to be a difference between the malignant and non-malignant exudate groups.

Table 4.2.2.1. A summary of serum laboratory parameter levels measured in 61 ascites patients by peritoneal effusion status (malignant ascites exudate MA, non-malignant ascites exudate NMA, ascites transudate TA) found to be significant at univariate analysis.

Laboratory parameter	Group	Mean	SD	Bonferroni adjusted p-value	Laboratory parameter	Group	Mean	SD	Bonferroni adjusted p-value
C3 (mg/L)	MA NMA TA	1713 1318 775.3	608.0 451.2 407.0	< 0.001[1]	**HAP#** (mg/dL)	MA NMA TA	341.7 285.5 106.0	128.4 152.7 63.1	< 0.001[1]
C4# (mg/L)	MA NMA TA	675.4 551.7 224.7	245.5 322.1 127.8	< 0.001[2]	**AAT#** (mg/dL)	MA NMA TA	476.8 357.1 291.0	145.9 125.7 80.6	< 0.001[3]
TNF-α (fmol/ml)	MA NMA TA	87.1 91.3 42.8	25.7 44.5 32.0	<0.001[1]	**AAG** (mg/dL)	MA NMA TA	167.0 147.2 88.8	53.7 73.0 54.2	<0.01[1]
IL-1a# (fmol/ml)	MA NMA TA	54.5 38.5 17.1	38.9 37.3 11.1	<0.001[1]					

[1] Grp MA vs Grp TA, Grp NMA vs Grp TA
[2] Grp MA vs Grp TA, Grp NMA vs Grp TA and Grp MA vs Grp NMA
[3] Grp MA vs Grp TA, Grp MA vs Grp NMA
[#] Log - transformed before 1-way ANOVA test

In the ascitic fluid, mean values of the following variables differed significantly between groups: the number of white blood cells (WBC), total protein (PROT), LDH, IL-1a, albumin (ALB), TNFa, ferritin, complement factor C3, complement factor C4, CER, AMG, HAP, AAG, AAT, TRF and interleukin - 8 (IL-8). Summaries are presented in **Table 4.2.2.2.** below. The groups between which the significant differences lie are provided by the SNK contrasts, from which there is an indication that, as in the serum measurements, it is usually the two exudate groups that differ

from the transudate group. For C4, albumin and AAT there also appears to be a difference between the malignant and non-malignant exudate groups.

Table 4.2.2.2. A summary of ascitic fluid laboratory parameter levels measured in 61 ascites patients by peritoneal effusion status (malignant exudates MA, non-malignant exudate NMA, transudate TA) found to be significant at univariate analysis.

Laboratory parameter	Group	Mean	SD	p-value	Laboratory parameter	Group	Mean	SD	p-value
WBC# (μL)	MA NMA TA	2588 1912 181.7	2500 2309 139.2	<0.001[1]	**IgA** (mg/dL)	MA NMA TA	153.5 173.7 66.4	93.7 109.0 58.4	0.01[1]
Protein (g/dL)	MA NMA TA	4.5 4.9 2.5	0.84 1.01 1.82	<0.001[1]	**IgG** (mg/dL)	MA NMA TA	778.5 1019.4 431.2	334.2 677.2 305.8	0.01[1]
Albumin (g/dL)	MA NMA TA	2.7 1.9 0.8	0.61 0.69 0.93	<0.001[2]	**CRP#** (mg/dL)	MA NMA TA	2.4 1.7 1.1	1.4 0.8 0.9	0.01[1]
LDH (IU/L)	MA NMA TA	461.3 458.9 89.1	287.0 379.0 27.5	<0.001[1]	**CER** (mg/dL)	MA NMA TA	23.4 23.4 9.9	9.9 8.0 4.9	<0.001[1]
FRT (ng/ml)	MA NMA TA	600.6 526.5 103.1	366.9 294.6 84.4	<0.001[1]	**TNF** (fmol/ml)	MA NMA TA	165.0 123.8 69.3	50.5 42.2 35.0	<0.001[1]
C3# (mg/L)	MA NMA TA	1350 804.6 417.0	537.9 369.4 359.0	<0.001[1]	**IL-1a#** (fmol/ml)	MA NMA TA	56.2 67.1 15.5	38.2 46.9 22.6	<0.001[1]
C4# (mg/L)	MA NMA TA	491.5 405.1 91.6	314.3 271.1 54.0	<0.001[2]	**IL-6#** (fmol/ml)	MA NMA TA	134.5 98.9 56.3	57.7 58.4 36.9	0.01[4]
AMG (mg/dL)	MA NMA TA	49.2 50.2 26.3	18.6 19.3 11.6	<0.001[1]	**IL-8** (pg/ml)	MA NMA TA	742.3 803.3 204.6	515.8 490.7 127.2	<0.001[1]
HAP# (mg/dL)	MA NMA TA	84.2 64.8 21.8	38.5 40.6 20.0	<0.001[1]	**AAT#** (mg/dL)	MA NMA TA	249.6 190.2 75.0	79.7 83.4 55.7	<0.001[2]
TRF (mg/dL)	MA NMA TA	120.2 95.4 57.1	48.3 40.4 32.3	<0.001[1]	**AAG** (mg/dL)	MA NMA TA	91.2 83.6 22.6	30.1 36.1 23.9	<0.001[1]

[1] Grp MA vs Grp TA, Grp NMA vs Grp TA
[2] Grp MA vs Grp TA, Grp NMA vs Grp TA and Grp MA vs Grp NMA
[3] Grp MA vs Grp TA, Grp MA vs Grp NMA
[4] Grp MA vs Grp TA
Log - transformed before 1-way ANOVA test

None of the single measurements resulted in complete discrimination between the three groups. The stepwise discriminant analysis resulted in the following five parameters being considered jointly to have the maximum discrimination power between the three groups: the ascitic fluid levels of PROT, LDH, TNF-α, C4 and HAP. The canonical discriminant functions used for classification are presented in **Table 4.2.2.3.** below.

Table 4.2.2.3. Canonical discriminant function coefficients derived from the stepwise discriminant analysis for three groups of ascites patients with peritoneal effusions*

	Canonical function	
	1	**2**
PROT	1.064	1.991
LDH	0.684	0.521
TNF	1.045	-1.944
C4	0.799	-0.344
HAP	0.678	-0.130
(Constant)	-16.341	6.119

* the coefficients are unstandardized and the data on natural logarithmic scale

For example, subject A with measurements of 3.30 g/dL, 244 IU/L, 276.98 fmol/L, 155 mg/L and 85.8 mg/dL for PRT, LDH, TNFa, C4 and HAP respectively would have the following scores:

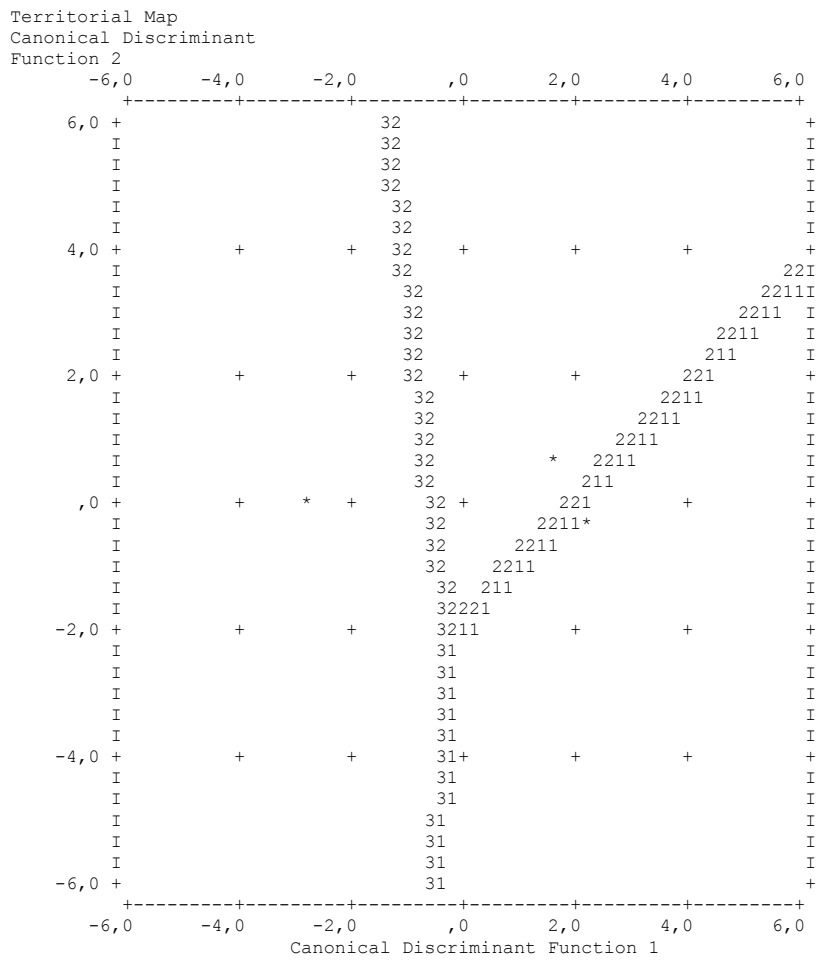
Canonical function 1: 1.064*ln(3.30) + 0.684*ln(244) + 1.045*ln(276.98) + 0.799*ln(155) + 0.678*ln(4.452) – 16.341 = 1.6158

Canonical function 2: 1.991*ln(3.30) + 0.521*ln(244) - 1.944*ln(276.98) - 0.344*ln(155) - 0.130*ln(4.452) + 6.119 = -1.8850

(where "ln" represents the natural logarithm)

As can be seen in the territorial map below (**Figure 4.2.2.1.)**, subject A would be classified as having a malignant exudate (MA).

Figure 4.2.2.2. below is a graphical scatter display of the individual canonical function scores in which the centroid of each group is also displayed. As can be seen from this scatter plot, differentiation between the transudates and the exudates is much clearer than differentiation between the two exudate groups. In the present investigation, a further step was taken by inclusion of the age and sex of the patients in the model before the stepwise variable selection procedure but the final model remained unchanged with only the five protein measurements being influential.



Symbols used in territorial map

Symbol	Group	Label
1	1	Malignantascites
2	2	Non-malignant exudates
3	3	Transudates
*		Indicates a group centroid

Figure 4.2.2.1. A territorial map of the canonical discriminant functions for group separation (1=MA,2= NMA,3= TA).

.

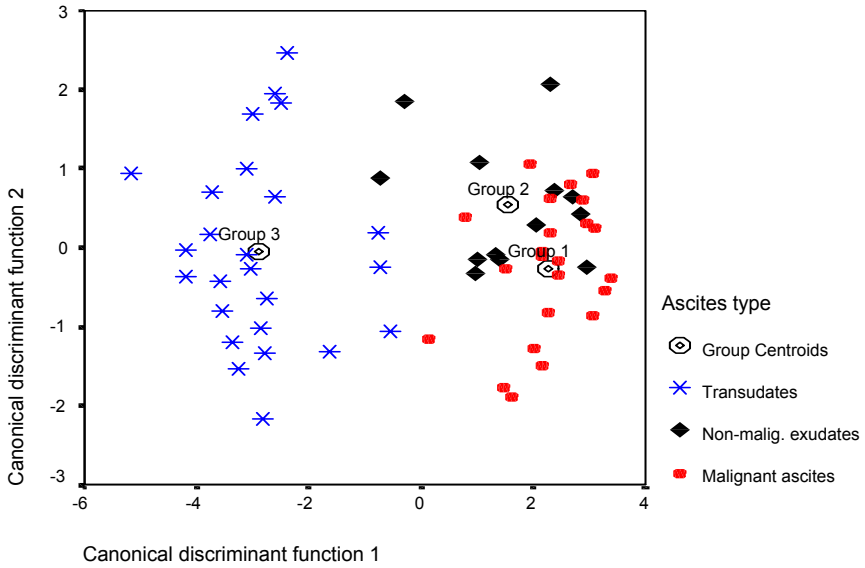


Figure 4.2.2.2. Canonical discriminant function scatter plot for the 3 groups of ascites patients

Overall, 89% of the cases were correctly classified, with 100% correct classification of the 25 transudates. In **Table 4.2.2.4.** it can be seen that the model correctly classified 19 of the 23 of patients in group MA, while the remaining four were classified as NMA.

Table 4.2.2.4 Classification table for discrimination between the three groups of ascitic patients

		Predicted Group Membership		
Actual Group	**No of Cases**	**MA**	**NMA**	**TA**
Group MA	23	19	4	0
	100%	*83%*	*17%*	*0%*
Group NMA	13	3	10	0
	100%	23%	*77%*	*0%*
Group TA	25	0	0	25
	100%	*0%*	0%	*100%*

Cross–validation was used to investigate the applicability of the final model to new data. With this method, each individual is classified based on the data of the other 60 cases. As can be seen in **Table 4.2.2.5.**, 69% of the patients were correctly classified under the cross – validation process.

Table 4.2.2.5 Classification table under cross validation

		Predicted Group Membership		
Actual Group	**No of Cases**	**MA**	**NMA**	**TA**
Group MA	23	16	7	0
	100%	*70%*	*30%*	*0%*
Group NMA	13	8	3	2
	100%	*62%*	*23%*	*15%*
Group TA	25	0	2	23
	100%	*0%*	*8%*	*92%*

4.3 Seroprevalence of viral markers in Crete

The crude prevalence rates of HBsAg and HCV in Cretan blood donors were estimated to be 0.40% and 0.38% respectively. Male blood donors had a higher prevalence of HBsAg compared with female blood donors (0.41% versus 0.28%, Z-statistic =5.28, $p<0.01$ and Cochran's test $p<0.01$). The estimated odds of HBSAg positivity were 98% higher for males (OR=1.98, 95% CI: 1.21 to 3.23). Exposure to HBV was detected in 8.8% of blood donors in Hania (5.9% of females and 9.2% of males, no evidence of a difference between sexes) and 9.1% in Heraklion (9.4% of males. 7.0% of females, again no evidence of a difference between sexes). Overall, the estimated odds of being exposed to HBV were 57% higher for males than females (OR:=1.56, 95% CI: 1.37, 1.80). Anti-HCV was detected in 0.38% of donors, with higher rates in Heraklion (0.52%) and Rethymnon (0.52%) than in Hania (0.23%). Significant differences between the three prefectures were also found for HBsAg levels, with Rethymnon having the lowest prevalence (0.27%).

In the hospital patients, the crude prevalences of HBsAg and anti-HCV were 2.66% and 4.75% respectively. A lower prevalence of HBsAg was detected in Rethymon patients than in thoses in the other two prefectures, both overall (1.46% cf 3.96% in Heraklion and 2.30% in Hania) and for males and females separately (2.03% in Rethymnon males c.f. 5.41% in Heraklion males and 2.80% in Hania males, $p<0.001$, and 0.94% in Rethymnon females c.f. 2.90% in Heraklion females and 1.73% in Hania females, $p<0.001$). In all three hospitals, the RR of positivity for HBsAg was significantly higher for males than females (5.4% vs. 2.9% in Heraklion hospital patients, RR=1.9, $p<0.001$, 2.8% vs. 1.7% in Hania hospital patients, RR=1.6, $p<0.001$, 2.0% vs. 0.9% in Rethymnon hospital patients, RR=2.2, $p<0.001$). The overall prevalence of HCV was 5.2% in Rethymnon, 2.4% in Hania and 6.6% in Heraklion. Similarly to HBsAg, there were significant differences between anti-HCV positivity rates in males and females in Heraklion (males 7.3%, females 6.0%, RR 1.23, $p<0.001$) and Hania (males 2 .7%, females 2.0%, RR 2.4, $p<0.01$).

In measurements taken from Cretan subjects from the general population, exposure to HBV was estimated at 29% and 25% in the urban and rural populations respectively

(no significant difference) whilst the crude carrier rates were estimated to be 2.7% and 0.8% respectively ($p<0.01$). The prevalence of anti-HCV was not found to differ significantly between urban and rural populations, being 4% in the urban population and 2% in the rural population. HCV-RNA was found in 50% of cases in the urban population and 25% in the rural population, although again this difference was not found to be statistically significant).

5. DISCUSSION

Patients with cirrhosis

Of our total cohort of 470 consecutive cirrhotis patients (who were 62% male), 320 (68%) had compensated cirrhosis at presentation and 144 (31%) had decompensated cirrhosis, whilst 6 (1%) were of unknown status. This is in contrast to a series of 1155 consecutive cirrhosis patients (who were 65% male), 63% of whom had features of decompensation at first presentation (D'Amico et al, 1986). The percentages by cirrhosis aetiology appear similar for the male cirrhotics in each cohort, with 33% of the males in the Italian cohort having alcohol abuse as aetiology and 14% of males being positive for HBsAg as compared to our 42% and 16% respectively. For females, however, there is a somewhat different picture, with 15% having alcohol abuse as cirrhosis cause and 6% being HBsAg-positive as contrasted with our 2% and 10% respectively. For our compensated cirrhosis patients, life expectancy is relatively long, with 67% of our patients surviving 7 years after diagnosis (**Table 4.1.1.6**) as compared to a 6-year survival rate of 54% in the Italian patients (D'Amico et al, 1986). The survival prognosis is poorer in both cohorts, however, for those patients in whom decompensation has already occurred, with 36% survival at 7 years in our group (**Table 4.1.1.12**) and a 6-year survival percentage of 21% in the Italian group.

In our patients, decompensation and survival rates were found to differ according to cirrhosis aetiology, with type C patients having lower risks of decompensation and death than the other aetiological groups (with RR of decompensation 0.58 compared to cryptogenics, **Table 4.1.1.3.**, and RR of death 0.45 compared to cryptogenics, **Table 4.1.1.7.**). The cirrhosis patients who have alcohol as aetiology have a higher risk of decompensation (RR 1.72, 95% C.I. 1.00 to 2.97) than other aetiological groups. There is no evidence, however, of a decreased survival time in alcoholic cirrhotics, whether they are diagnosed as compensated or whether they present with decompensation.

Overall, 49% of our compensated cirrhosis patients remain decompensation-free 5 years after diagnosis (**Table 4.1.1.2**). The predicted rates of hepatic decompensation in the present study may be higher in our type B and C patients than in those of similar studies for patients with virus-related cirrhosis. The percentage of HBsAg-

positive patients who remain decompensation-free 5 years after diagnosis in the present study is only 30% (**Table 4.1.1.2**) whereas a multicentre retrospective study undertaken within the European concerted action on viral hepatitis (EUROHEP) involving nine hospitals in Western Europe estimated the percentage of HBV cirrhotics decompensating 5 years after diagnosis to be 23% (Fattovich et al, 1995). The percentages of our HCV patients estimated to remain decompensation-free at 3 and 5 years after diagnosis are 79% and 65% respectively (**Table 4.1.1.2**) whereas a EUROHEP study found cumulative percentages of patients developing decompensation at 3 years and 5 years after diagnosis to be only 12% and 18% respectively (Fattovich et al, 1997). A further European study of 103 compensated cirrhosis type C patients in France found a 4-year risk of decompensation of 20% (Serfaty et al, 1998). Both in the EUROHEP cohort and in the French cohort, however, the majority of patients had received treatment during the follow-up period: 226 of the 384 patients in the EUROHEP cohort (59%) and 57% of the French cohort, the latter having been treated exclusively with interferon (IFN). Absence of IFN therapy has been found to be a predictor of decompensation and death (and HCC) in 103 HCV cirrhotics in France (Serfaty et al, 1998). Only 32% of our compensated HCV cirrhotics received treatment (***Plaquenil***) and there is no firm evidence that this treatment is a predictor of decompensation or death.

The predicted survival rates for compensated HCV patients in our study appear comparable to those of other studies. The percentages of our compensated HCV patients estimated to survive 3- and 5- years after diagnosis are 94% and 88% respectively (**Table 4.1.1.6.**) and corresponding percentages from a EUROHEP study are 96% and 91% (Fattovich et al, 1997). For compensated HBV patients, however, the predicted survival rates appear somewhat lower in our study than in the corresponding EUROHEP study. The percentages of our compensated HBV patients estimated to survive 5- years after diagnosis are 64% (**Table 4.1.1.6.**) as compared to 84% in both the EUROHEP cohort (Realdi et al, 1994) and another group of HBsAg-positive compensated cirrhotics (de Jongh et al, 1992). The EUROHEP investigators suggested, however, that the high predicted survival rates in their cohort may be due to differences in patient characteristics between study cohorts, particularly as their cohort had a low rate of splenomegaly (29%) and hepatic stigmata (28%), suggesting

a relatively early stage of the disease. These indicators were not available for our retrospective analyses.

The percentages of our decompensated HCV patients estimated to survive 3 and 5 years after presentation are 69% and 50% respectively (**Table 4.1.1.12.**). These survival rates appear very similar to those of the EUROHEP cohort of type C cirrhosis patients, estimated after the first appearance of decompensation, at 5 years to be 50% (Fattovich et al, 1997). The percentages of our decompensated HBV patients estimated to survive 3 and 5 years after presentation are 34% and 26% respectively (**Table 4.1.1.12.**). The 5-year survival of a group of Dutch type B decompensated cirrhotics was only 14% (de Jongh et al, 1992) whereas the corresponding EUROHEP cohort had survival of 50% after decompensation (Fattovich et al, 1995). The survival chances of our cirrhotics who present with ascites appear to be high, as a recent study of 216 Spanish cirrhotics with ascites found a 27% chance of survival after 5 years (Fernandez-Esparrach et al, 2001), as compared to our 52% survival percentage (with 95% C.I. 39% to 64%) for the equivalent group (**Table 4.1.1.12**).

The above findings indicate that in our cirrhosis patients, whilst overall survival rates for both compensated and decompensated cirrhotics are likely to be similar to those of other European groups of cirrhotics, rates of decompensation in our compensated types B and C cirrhosis groups may well be higher than in other European cohorts. One explanation for the increased rates of decompensation in our viral cirrhosis groups may be that the mean ages at diagnosis were 63 years (s.e. 2.1) and 64 years (s.e. 0.7) for HBV and HCV patients respectively as compared to means of 44 years (range 17-74) and 54 years (s.d. 5 years) in the corresponding EUROHEP cohorts (Realdi et al, 1994, Fattovich et al, 1997). Age has been found, both in the present study and from the EUROHEP Cox models, to be a significant prognostic factor for both decompensation and survival (Realdi et al, 1994, Fattovich et al, 1997, Fattovich et al, 2000). Fattovich et al (1997) found that the probability of survival in type C cirrhosis depended on the presumed source of infection, with better chances of survival in patients with a history of intravenous drug abuse or blood transfusion than other sources (such as a family member with chronic liver disease). When considering the survival times of our decompensated patients, it should be noted that they were followed from the time of their presentation to the clinic which was not necessarily

the same as the time of diagnosis (29% presented at the clinic some time after the initial diagnosis) whereas the EUROHEP estimates were taken using the time of initial decompensation in the cohort of compensated patients as the starting point. It may well be, therefore, that our decompensated cirrhosis patients have longer survival times on average than other European groups when considered from the time of diagnosis.

Although information on the survival rates of cirrhosis patients with alcoholism as aetiology is scarce, liver disease is known to progress more rapidly amongst persons with joint alcoholic liver disease and HCV infection than those with HCV alone (WHO fact sheet 164, 2000). Our findings indicate that this holds even when the disease has already progressed to the cirrhotic stage (**Tables 4.1.1.1 and 4.1.1.5**), with deterioration also being more rapid than in those subjects with only alcoholic liver disease.

The rate of HCC incidence in Cretan cirrhotic patients was estimated to be 2.3 per 100 person-years. In other studies, incidences range from 1.5 per 100 person-years in 349 HbsAg positive compensated cirrhotics (Fattovich et al, 1995) to 6.4 per 100 person-years in male compensated cirrhotics who were both HBsAg positive and anti-HCV positive (Chiaramonte et al, 1999). The overall cumulative 3- year HCC incidence rate (obtained using estimates of the cumulative hazard) in our cirrhosis patients was estimated to be 8% (se 1.5%). In 240 patients with cirrhosis diagnosed at enrollment to a study undertaken in Osaka, Japan, the estimated cumulative risk (estimated using Kaplan-Meier methodology, as in the present estimates) was 12.5%, with se 2.5% (Tsukuma et al, 1993).

In our cirrhosis type B patients the cumulative 3- and 5-year HCC incidence rates were 20% and 27% respectively whilst in our type C patients the corresponding percentages were only 7% and 9% respectively. These results are in the opposite direction to those obtained in a study of 259 Italian compensated cirrhotics with cumulative 5-year HCC appearance rates in HBsAg-positive cirrhotics at 10% and in HCV-positive cirrhotics at 21% (Chiramonte et al, 1999) and to those obtained in a follow-up study of 795 cirrhosis patients in Japan, with HCC appearance rates at the fifth year at 14% in the HBsAg positive patients and at 22% in the HCV positive

patients (Ikeda et al, 1993). The cumulative incidence rates found for HCV patients in our study were similar to the corresponding EUROHEP percentages: 4% at 3 years and 7% at 5 years (Fattovich et al, 1997). For the HBV patients, the corresponding EUROHEP estimates were much lower: 3% at 3 years and 6% at 5 years (Fattovich et al, 1997). One possible explanation for these differences is that in our cohort, only 55 HBV patients were included in the analyses with 9 cases of HCC appearance during follow-up.

A somewhat surprising result was that the incidence rate in our compensated Cretan cirrhotics was higher than in the patients with decompensated cirrhosis (2.5 per 100 person-years and 1.5 per 100 person-years respectively). However, deaths from liver disease are likely to override HCC incidence for given time intervals involving advanced stages of decompensated cirrhosis. A "survival bias" may be introduced as a result of reduced progression rate to HCC compared to earlier time points, due to deaths as a result of decompensation (Chiaramonte et al, 1999). These deaths are treated as censored observations in the estimation of HCC incidence rates, as in similar studies (e.g. Ikeda et al, 1993). This provides an explanation for the observation that the estimated incidence rate is higher in the compensated as opposed to the decompensated cirrhotics. An alternative explanation is that the estimates are approximate, particularly given the small number of decompensated cirrhotics. Another plausible explanation for the unexpected incidence rates is that the estimated rates are crude rates. Age-adjusted rates may provide more reliable estimates, although these are not estimable in the present study, given the small numbers involved.

In the compensated cirrhosis group, the proportion of type C patients was higher than in the decompensated patients (48% and 16% respectively) whilst the proportion of those with alcohol as aetiology was lower (18% c.f. 49%). Cirrhosis aetiology is thought to have an effect on the risk of HCC, and this is also indicated by the present findings with higher incidences in the HBV as compared to the HCV cirrhotics. The differing HCC incidence rates may be an explanation for the differing survival rates found in the aetiological groups. Some previous studies have found an association between the cause of cirrhosis and the number of HCC nodules (Fasani et al, 1999) whilst others have found HCC to grow more aggressively in patients with HBV than

those with HCV (Okuda et al, 1984, Shijo et al, 1991, cited in Fasani et al, 1999). A higher prevalence of HCC in cirrhosis patients with multiple aetiologies than in HCV carriers has also been reported (Fasani et al, 1999). In a study of 917 outpatients with chronic hepatitis or compensated liver cirrhosis, each of the serum markers for hepatitis virus (HBsAg, anti-HBC in high titre and anti-HCV) was significantly associated with the risk of liver cancer, as was the diagnosis of liver cirrhosis at enrollment to the study. The amount of alcohol consumed per day (>80g ethanol versus <80g ethanol) did not, however, have an effect on the risk of liver cancer (Tsukuma et al, 1993).

The study of the survival of Cretan cirrhotic patients has all the inherent limitations of a retrospective analysis. One limitation of the study is that the severity of cirrhosis at diagnosis was not recorded. In fact, there were only demographic and aetiological prognostic factors available for inclusion in the survival and decompensation models. Therefore, clinically useful inferences are limited. For example, it is well known that the presence of HCC or decompensation are stronger risk factors for early death than cirrhosis aetiology, age or sex and determination of the extent to which the former preside over the other risk factors is not likely to be of primary importance. The availability of measurements such as that of serum albumin may be of high inferential use, this measure having been associated with prognosis in cirrhosis due to alcohol, hepatitis B and cryptogenic causes (Gines et al, 1987, cited in Bonis et al, 1999) and also cirrhosis due to hepatitis C (Fattovich et al, 1997). Fattovich et al (2000) found the serum γ-globulin level to be a significant prognostic factor in the probability of decompensation and survival of HBV cirrhotics and offer the explanation that the levels of this biochemical variable reflect the degree of alteration of the hepatic circulation in the cirrhotic liver. Available α-fetoprotein (AFP) measurements at presentation may also be beneficial, particularly as it has been suggested that patients with levels of greater than 20μg/l may have undiagnosed HCC (Colombo et al, 1991). The limited value of the prognostic information available for the models presented is reflected in the short-term predictive accuracy of the multivariable Cox model, as compared to the simple Kaplan-Meier estimates, as depicted in the time-to-decompensation model, for which R^2 falls to 0 after 5 years (**Section 4.1.1.**).

It may also be beneficial to have genotype assays for HCV. The most common genotypes in European patients with cirrhosis type C are genotypes 1b and to a lesser extent, type 2 (Fattovich, 2001). In cirrhotics, HCV type 1b is not associated with a greater risk for HCC compared to other genotypes but patients with this genotype have a threefold increase in the risk of decompensation (Fattovich, 2001). Some previous reports indicate that genotype 1b may be over represented amongst HCV patients with cirrhosis and HCC (Nousbaum et al, 1995, Hatzakis et al, 1996, Zein et al, 1996 cited in Bonis et al, 1999). A separate study, however, found no such association (Bonis et al, 1999). It is thought that HCV-infected subjects with higher amounts of HCV RNA in their serum (i.e. higher viral loads) are more resistant to interferon therapy (Di Bisceglie, 1998). It would be interesting to measure the viral loads in our HCV treated patients. It is also believed that sustained responses to interferon are more likely to be obtained in those with a shorter duration ot the disease, milder histological features, genotypes other than 1b and limited quasispecies diversity (Dienstag, 1997).

A co-infection with B and C viruses has been found in certain studies to produce more accelerated disease of the liver indicating possible synergistic effects of each infecting genotype, rather than additive (Roudot-Thoraval et al, 1997, cited in Bonis et al, 1999, Benvegnu et al, 1994, cited in Fasani et al, 1999). In the present study, however, the number of subjects with dual infection was too small (2 males and 1 female presenting with compensated cirrhosis due to dual B and C infection, 2 males presenting with compensated cirrhosis due to dual B and Δ infection and 1 male presenting with decompensated cirrhosis due to B,C and Δ infection) to allow inferences to be drawn. In a recent EUROHEP investigation, a 20% prevalence of anti-HDV was found in 200 HBsAg positive compensated cirrhotics, a rate said to parallel those of previous European studies (Fattovich et al, 2000). A somewhat higher risk of HCC and mortality was found in the HDV-infected HBV cirrhotic patients, although the mortality results did not quite reach statistical significance.

The presence of hepatitis B e antigen (HBeAg, a marker of viral replication) in HBsAg-positive patients may also be of prognostic significance. A EUROHEP study of survival of compensated type B cirrhotics found HBeAg status to be one of six significant prognostic factors for survival (Realdi et al, 1994). A further EUROHEP

study of HBV cirrhotics found that HBeAg positivity at entry was the only significant prognostic factor (in a Cox model) for HBsAg loss at a later date and that the loss of HBsAg was associated with a low risk of developing HCC and long survival (Fattovich et al, 1998). In the same study, the yearly incidence of HBsAg loss in untreated patients with compensated cirrhosis type B was estimated to be 0.8% during the first five years of follow-up. A study of a large cohort Dutch type B cirrhotics also supported the idea of HBeAg status being an important prognostic indicator for survival, with a change in HBeAg status during follow-up resulting in a 55% decrease in mortality rate (de Jongh et al, 1992). The EUROHEP investigators also studied the effect of **interferon –alpha** and found a significantly higher probability of HBsAg loss in the treated group as compared to the untreated controls (Fattovich et al, 1998). Another interesting finding was that in the interferon-alpha treated group, HBsAg loss occurred only if HBeAg had been present at entry.

Another limitation of the present investigation is that the results may not be strictly representative of the natural history of cirrhosis, given that a minority of the type C patients received treatment, albeit without significant survival implications (details in **Section 4.1.2.**). A separate issue is that of the effects of possible pre-existing end-stage liver disease. Bonis et al (1999), who developed predictive models for the development of HCC, liver failure or liver transplantation in patients presenting with chronic hepatitis C, state that "those who developed a primary end-point within 4 months of initial evaluation were excluded [from the analysis] since it is likely that end-stage liver disease already existed". In the present study, a similar exclusion criterion was enforced in the case of HCC incidence (with exclusion of those for whom HCC occurred within one month of initial evaluation, details in **Section 3.1**). Further models could be developed for our data using similar exclusion criteria for decompensation and liver failure.

PBC data

The results of our study on UDCA-treated PBC patients indicate that in our PBC population there is a beneficial effect of UDCA on long-term survival. The present study, however, has the inherent limitations of any trial that is not randomized and controlled, particularly as it has a long accrual period. The Mayo model patients were, however, also accrued over a long time period (January 1974 to May 1984, Grambsch

et al, 1989), albeit without being administered treatment for PBC. Our study has the statistical advantage over other similar studies that there was a single clear end-point in the analysis which did not involve referral for liver transplantation. The criteria for referral for liver transplantation may vary between different countries and locations, resulting in differences in possible interpretations of UDCA benefits, depending on the particular study (Goulis et al, 1999). As liver transplantation was not an option for the Cretan PBC patients, this was not a potential source of bias in the estimation of observed survival per se in our data.

The Mayo model seems to stratify the patients well according to risk of death in that the risk score was found to be the most important predictor of survival status using Cox regression analysis, both considered singly and in combination with other prognostic variables. Dichotomizing the patients based on their risk scores resulted in two groups of patients with those in the low risk score group having significantly better chances of survival than their high risk score counterparts. When the five variables considered in the Cox regression-derived Mayo model, however, were considered in a similar Cox regression analysis for our data, the only significant predictors of survival time were found to be initial age and log(albumin) concentration. Our Cox analysis results must be interpreted with care, however, due to the strong evidence of lack of stability of the model coefficients.

There are certain important points to be considered in the quantitative application of Mayo natural history model predictions to compare survival of untreated PBC patients to survival of UDCA-treated patients. Firstly, the Mayo model has been applied to a Greek population whose baseline characteristics may not be similar to those of the original population. In the Ludwig staging classification, there appear to be a higher proportion of Greek PBC patients initially at Ludwig stage I or II (58% in total) than in the 418 patients from which the Mayo model coefficients were derived (28% in total at stage I or II, Dickson et al, 1989). In the Mayo model, the data were collected at randomization to a clinical trial (the starting date being the date of entry to the trial) whereas in Crete, the starting point was the time of disease diagnosis. The differences observed in the two patient groups tie in with the measurements being taken at diagnosis as opposed to the date of enrollment into the study. It is notable that the initial average age of the patients was higher in the Cretan patients, however, with

median 59 years, than in the Mayo model patients (median 50 years). The mean risk score of our patients was 4.83 (sd 1.21, range 2.32 to 8.93), a value very similar to that of the Mayo model mean of 5.07 (range 2.78 to 10.17) for the combined Mayo data set (418 patients, Markus et al, 1989). Also, the Mayo natural history model was developed over a decade before the present data were analysed. During this time period there may have been improvements in baseline care other than the treatment effects. A double-blind randomized study of UDCA-treated versus placebo groups of PBC patients, comparing to Mayo model predictions have shown there to be an inflated effect of treatment based on the Mayo model (Lindor et al, 1996). Based on the above observations it is unclear in which direction, if any, the resulting survival bias may be, although it does appear that it should not be taken for granted that the baseline hazard in the Cretan PBC patients is the same as that of the Mayo model population. A second point to consider is the use of the one sample log-rank test procedure in comparisons with a hypothetical control group is not ideal, as the mean survival function is random and not fixed as assumed in applying the test (Dickson et al, 1989). This test is, however, the one of choice in many such comparisons and has been widely applied in making Mayo model comparisons using PBC patients (e.g. W. Ray Kim et al, 2000, Poupon RE et al, 1999, Krzeski et al, 1999, Markus et al, 1989).

It is hoped that in the near future it will be possible for liver transplantations to be undertaken in Crete. The survival probabilities of our PBC patients could be used to assess when transplantation could take place. It has previously been suggested that transplantation should be considered when the estimated 6-month survival probability drops below 80% (Christensen et al, 1993). As there was a high degree of censoring (84% censored values) in our data set, it was not possible to develop a prognostic model containing all known information. In fact, it has been suggested that the number of events (deaths) per variable (EPV) considered for inclusion in a PH model should be at least 10, as with a smaller number the parameter estimates in PH models have been found to be unreliable (Peduzzi et al, 1995, Peduzzi et al, 1996 cited in Altman & Royston, 2000). Using the present data, the EPV for a two-variable model is 8.5 and for a three-variable model it is 5.7.

It is a well-documented observation that in untreated PBC patients, elevated levels of serum bilirubin are an independent predictor of a poor prognosis (Wiesner, 1998,

Shapiro et al, 1979, Dickson et al, 1989). This has also found to be true for UDCA-treated patients in a trial in which comparisons were made between patients with 'normalised serum bilirubin level' (i.e. <17 μmol/L, continuing on a consecutive measurement) and those without normalized levels (Bonnand et al, 1999). In a recent study of Polish PBC patients, bilirubin was found to be the most important predictor of prognosis, whether or not there had been prior treatment with UDCA (Krzeski et al, 1999). The importance of bilirubin is also reflected in the standard prognostic models (Mayo, European, Yale, Oslo) in which the level of bilirubin is the most heavily weighted variable (Wiesner, 1998). It is interesting that in our UDCA-treated patients, bilirubin levels did not appear to be a significant predictor of survival (neither at univariate nor multivariate analysis). This observation may reflect relatively early diagnosis of the patients in the study, many of whom were asymptomatic for PBC and were diagnosed as a result of routine blood tests. Perhaps these results would be different if a time-dependent model was considered, with repeated measurements of bilirubin over time. Recently, a new Mayo model has been developed which incorporates repeated measurements, leading to higher accuracy and precision in the two years following the patient's last visit (Wiesner, 1998).

A separate issue is that of the extent of development of complications of liver disease such as ascites or gastrointestinal bleeding, which may be considered as clinically meaningful surrogates for survival. In the present cohort, the subjects in whom ascites or variceal bleeding occurred were given standard treatment (paracentesis or diuretics for ascites and sclerotherapy or band ligation for variceal bleeding).

Natural history of HCC

From the results of the survival analysis of 73 consecutively diagnosed HCC patients, it could be cautiously inferred that the situation with regard to the relationship between HCC and the viral markers HBV and HCV in Cretan cirrhotics is in contrast to that present in mainland Greece, with HCC being associated with 54% HCV and only 26% HBsAg, our results approaching those of Japan, Italy and Spain (further details having been provided in **Section 1.4**). This may be due to the fact that the prevalence of HBsAg on Crete is lower than in mainland Greece (with an estimated overall prevalence in blood donors on Crete of 0.40%, Koulentaki et al, 1999, as compared to 0.84% in sporadic donors in Greece, Kyriakis et al, 2000), and resembles

the situation in Spain and Japan whilst anti-HCV positivity has been found to be higher in both rural and urban Cretan populations than in the population of mainland Greece, both in the present study and in previous investigations (Lionis et al 1997a, Lionis et al 1997b, Fragiadakis, 1996). The finding that HCC is associated more frequently with HCV than HBsAg may appear at first glance to be in contrast to the finding that in the Cretan cirrhosis cohort, the incidence of HCC is at much higher rates in the HBV than the HCV group. One explanation is that, given the higher prevalence of chronic HCV than HBV in Crete (and assuming similar rates of progression of both HBV and HCV to cirrhosis), the higher incidence rates of HCC found in our HBV cirrhotics may be more than compensated for by the difference in prevalences in the general population, resulting in more HCC patients being type C than type B. In fact, hepatitis B is believed to have a higher cancer potential than hepatitis C, a fact consistent with the finding of higher HCC incidence rates in type B Cretan cirrhotics. A further postulation is that the ***plaquenil*** treatment may have had a positive effect with respect to decreasing the incidence of HCC, on the HCV cirrhotics to whom it was administered, whereas none of the type B cirrhotics were administered any form of treatment.

It is noteworthy that of the patients diagnosed as having HCC during the follow-up period (1992-1996), small hepatocellular tumours (Okuda stage I) were rarely identified. In the publication of the analysis of these data, it was stated that it would be essential to create a surveillance programme of patients with HCV or HBV chronic liver disease (Kouroumalis et al, 1997). This surveillance programme has not yet been implemented.

Treatment of HCC

The first published study of the effect of treating inoperable hepatocellular carcinoma with somatostatin analogues and demonstrating improved survival with subcutaneous ***octreotide*** administration in such patients came from our clinic (Kouroumalis et al, 1997) and the survival analyses which were undertaken have been presented in the present study. Additional evidence has been reported from Austria as a case report (Raderer et al, 1999) whilst in Germany another a trial along similar lines is in progress, with the administration of long acting octreotide in patients with inoperable hepatocellular carcinoma (Allgaier et al 2000). Hepatocellular carcinoma has a

variable and heterogeneous clinical presentation and course, making the design of controlled trials for assessment of a new treatment modality extremely difficult (D Shouval, 1998). These difficulties were overcome in the case of our ***octreotide*** study, with our results being referred to as 'promising' and 'deserve to be explored further' (D Shouval, 1998). Another recent published study assessing the survival of Greek HCC patients was the study of ***tamoxifen*** treatment (Manesis et al, 1995). The estimated median survival times and survival rates were lower than in both our somatostatin analogue studies, with 22% survival after 12 months (as contrasted with 56% and 61% for our octreotide- and long-acting somatostatin-treated patients respectively). In both Greek studies, patients with tumours judged to be suitable for surgery at diagnosis were excluded.

In the patients involved in the octreotide study, two distinct groupings of somatostatin receptors were detected, in terms of their concentrations in the liver tissue (fmol/mg protein), unrelated to the underlying liver pathology (Kouroumalis et al, 1998). One possible explanation is a heterogenous distribution of somatostatin receptors in the tumoral tissue. This has been reported in certain adenocarcinomas (Reubi et al, 1990), pituitary adenomas (Greenman & Melmed, 1994) and carcinoid tumours (Reubi et al, 1994). In the long-acting somatostatin patient group, there appeared to be two distinct patterns: some patients remained remarkably stable for months with a biochemical improvement and in some instances with a recession of tumour size whereas other patients followed a relatively rapid deterioration leading to an early death. It is speculated that this could be related to the presence or absence of somatostatin receptors in the tumoural tissue (Dr D Samonakis, personal communication), although it was not possible to assess this quantitatively. With regard to the statistical analysis, it would be very interesting if the appropriate variable (e.g. receptor density) could be included in future prognostic models.

Ascites patients: differentiation between malignant and non-malignant ascites

Differentiation between malignancy-related and non-malignant ascites remains a difficult task. The results of the recursive partitioning models fitted in the present study indicate that, of the biological parameters considered, the most important factor in distinguishing patients with non-malignant (cirrhotic) peritoneal effusions from those with malignant ascites is the ascitic fluid:serum albumin ratio. These findings

are similar to those of a previous study, which indicated the importance of the serum:ascitic fluid albumin concentration gradient (Lee et al, 1992). There was also an indication that the ascitic fluid: serum interleukin-1a ratio may be of importance, although the split involved separation of only five subjects (three cirrhotics and two with malignant neoplasms) from the remaining twenty-five patients and when more stringent modelling conditions were used, the result was a misclassification rate of only 6% using the albumin ratio alone (as expected). Given the high observed correlations between the variables and high empirical sensitivities and specificities, the presence of other measurements possibly of similar importance to interleukin-1a (such as LDH and ferritin) must be considered a possibility. It is known that the presence of masking may complicate covariate evaluation in tree models (pg 102, Segal, 1998). Therefore, the model fitted in the present setting should be seen only as being indicative of the biochemical parameters that may be important in distinguishing cirrhosis from malignancy when based solely on the biochemical measurements. The small number of patients in the study increases the uncertainty of the validity of the specific model for patients other than those in the present trial, given the relatively large number of biochemical variables. The present investigation should be regarded as being a preliminary analysis, requiring validation in a prospective setting. Despite the aforementioned shortcomings, however, it is believed that this illustrates a simple and potentially very accurate model that renders the present study important as a basis for further research.

AAT in ascitic fluid has been reported to be a 95% specific and sensitive marker to separate malignant and non-malignant ascites (Villamil, 1990). The present results partially support these findings at a univariate level when using ascitic fluid to serum ratios, with a lower sensitivity (70%) and specificity (65%, see **Table 4.2.1.1., Section 4.2.1.**). AAT, LDH and ferritin have also recently been reported to have high sensitivity but low specificity in separating malignant from non-malignant pleural fluid (Alexandrakis et al, 1997). LDH levels in peritoneal fluid of ovarian cancers have also been used as a marker of diagnosis (Schneider et al, 1997). In the present study, high sensitivities and also high specificities were obtained for both LDH and ferritin. Complement measurements, reported to be useful tests in malignant ascites (Wang et al, 1997), offer no advantage over simple albumin measurements according to our experience.

An issue for consideration is the heterogeneity of the cancer patients in terms of tumour location. Among the 27 ascities patients with carcinomas, 12 had carcinomas of the ovary. The serum markers of TPS and CA-125 have been reported to be of additive value for the identification of epithelial ovarian neoplasms (Schneider et al, 1997). A higher proportion of women than men in the sample were cancer patients (22 females, 15 males) whereas the cirrhosis group consisted in the main of men (8 females, 15 males). A further issue is that, in the absence of further examinations, patients with malignant ascites cannot be further classified as to the presence or absence of cirrhosis. The statistical analysis aimed to discriminate between patients with malignant ascites and those ascitics with cirrhosis and without malignancy, using a restricted set of biochemical parameters. The ascitic patients with cirrhosis and without malignancy may have different biochemical characteristics to the general population of patients with ascites but without malignancy.

It is interesting that in the simultaneous discrimination between the three groups of patients (malignant exudates, non-malignant exudates, transudates), the concentration of albumin in the ascitic fluid was not one of the variables considered significant in the final multivariate model. This does not mean that albumin is not a useful discriminatory variable in this case, but rather that during the modelling procedure, when considering the variables jointly it was not necessary to include the albumin measurement to achieve a high discriminatory power for the three groups. Judging by the results of the univariate analysis, however, the importance of the ascitic fluid albumin concentration as a single predictor in distinguishing between the three groups is very clear (see **Table 4.2.2.2, Section 4.2.2.**).

Seroprevalence of viral markers

The viral marker seroprevalence estimates vary greatly between the populations considered (blood donors, high-risk hospital patients, community-based urban and rural populations). One common feature, however, appears to be their geographical distribution. Lower prevalence rates of HBsAg were found in Rethymnon compared to the other two counties, both in the blood donors and in the hospital patients. Similarly, lower rates of anti-HCV were found in Hania compared to the other two counties. Also, in general, lower rates of both anti-HCV and HBsAg positivity were

found in females as compared to males at each location. In our blood donor population, there appears to be a high exposure (8.8%) to and low carrier rate (0.40%) of HBV in comparison to other European countries. For example, Sweden has been reported to have an overall exposure rate to HBV of 3.6% and a carrier prevalence rate of 0.6% (Banke et al, 1971, Hansson et al, 1975, Iwarsson et al, 1972, cited in Koulentaki et al, 1999). A recent estimate of national HBsAg prevalence rates in non-regular donors was 0.84% (Kyriakis et al, 2000) whilst a study of blood donors in northwest Greece reported a HBsAg prevalence of 0.85% (Zervou et al, 2001). This is lower than previous HBsAg blood donor rates in selected Greek groups e.g. 4.9% in 6708 Hellenic Air recruits tested in 1971 (Vissoulis et al, 1972 cited in Kyriakis et al, 2000). It has been postulated that there has been a fall in the HBsAg seroprevalence in the general population in Greece in recent years (Kyriakis et al, 2000). The Cretan blood donor HBsAg seroprevalence estimates at 0.40% are seen to be even lower than the general Greek donor estimates mentioned above.

It is widely known that the estimated prevalence of viral markers in blood donors is likely not to reflect the prevalence in the general population. In fact, estimates obtained using blood donor populations are likely to be lower than those in the general population: in the U.S. the prevalence of HCV infection among volunteer blood donors in 1900 was only one third that of the general population (0.6% and 1.8% respectively, Wasley & Alter, 2000). This finding is confirmed in our study, where the general population HCV prevalence estimate of 3.0% is about eight times that of the blood donor population estimate whilst the HBV estimate of 1.7% is about 4.3 times that of the corresponding blood donor population estimate. As blood donors are frequently used in such surveys, however, it is possible to compare prevalence rates between countries using estimates for blood donor populations.

The crude anti-HCV blood donor prevalence rate at 0.38% appears to be at a similar or lower rate than in blood donor populations of other southern European countries. For example, in Spain anti-HCV prevalence rates of 0.93% and 1.12% have recently been reported (Munoz-Gomez et al, 1996, Salmeron et al, 1996) whilst in Italy a prevalence estimate obtained from 4614 blood donors of 0.3% has been reported (Meliconi et al, 1996). On the other hand, estimates of HCV seroprevalence rates in Crete obtained using both our community-based estimates (4.0% in the urban sample

and 2% in the rural sample) and those of a study carried out in primary health care (PHC) centres throughout Greece have been found to be higher than those obtained in other regions of Greece. In the latter study, the HCV carrier rate was found to be higher in Cretan centres (4.8%) than in those found in other regions in Greece (Lionis et al, 2000). The lowest prevalence in the PHC study was found in Macedonia, at 2.1%. The high seroprevalence estimates of anti-HCV in Crete compared to other regions of Greece are also in accordance with the findings of a previous study undertaken in rural Crete, with 10.9% of the 257 subjects tested when visiting the local health centre being found positive for anti-HCV and 3% of 164 subjects tested from surrounding villages found to be positive for anti-HCV (Lionis et al, 1997b). The corresponding prevalences of HBsAg were 1.2% and 0 respectively. Using the community-based HBsAg and anti-HCV positivity estimates obtained in our study and the estimates of previous studies undertaken in Crete, it can be inferred that the prevalence of HCV is likely to be higher than that of HBV in the Cretan population.

One major setback in all three epidemiological surveys undertaken in the present study is the lack of detailed breakdown by demographic factors such as age. Age is very likely to be a confounding factor (as, for example, age has a negative correlation with intravenous drug use). Young males are known to have higher rates of participation in risk activity which can result in exposure to blood borne viruses such as hepatitis B or C and males are also more likely if infected with hepatitis B to go on to be chronic carriers. In estimation of the HBV and HCV marker prevalence rates, the ages of individual subjects were not available in any of the groups considered although retrospective data were obtained at a later date on the distribution of age groups. It is mandatory that all males undertaking national service donate blood. Therefore, it is likely that the average age of the blood donors is lower than that in the general population. In addition, it would have been useful to have been provided with knowledge of the permanent addresses of the blood donors as a substantial proportion of those doing their National Service in Crete may not be permanent residents of Crete, so the rates may not be in fact truly representative of Cretan blood donor rates. It would also have been useful to have details for each subject on their status: regular or sporadic donor, and if sporadic, whether military recruit or family donor. National blood supply in recent years have been found to be 53% from directed family donors, 37% from regular donors, 6% from Hellenic Armed forces donors and 4% from Swiss

Red Cross donors (Hellenic Blood Transfusion Service, 1991-96, cited in Kyriakis et al, 2000). For nosocomial patients, similar problems of age bias are likely to be encountered, this time in the opposite direction i.e. older subjects on average than in the general population. The estimates obtained using the sample from urban and rural areas of Crete are likely to be the most representative of the seroprevalence rates in the general population in Crete.

General limitations of the statistical analyses

The statistical analyses that were undertaken can only provide an indication of the true state of affairs as they were limited by many factors. Two main limitations are presented below:

1) Many of the data analyses involved relatively small numbers of patients. This was particularly true for the multivariate analyses undertaken for the patients with ascites. It is known that with small numbers of patients, there is a low signal-to-noise ratio with an increased risk of selecting unimportant variables and failing to include important ones (Altman & Royston, 2000). Therefore, the results should be treated only as indicative. For the cirrhosis data, although associations were found between prognostic factors such as sex and cirrhosis aetiology (**Table 3.1.3.**), it was not possible to test for interactions in all the survival models, as the numbers in certain cells were too small. The presence of interaction terms, such as the effect of cirrhosis aetiology on the hazard of death being different in the patients of each sex, would affect interpretation of the results. Another example of the effect of small numbers is the lack of stability of the Cox survival model for the compensated cirrhotics, in which omission of two observations caused a large difference in the regression coefficient (and hence the hazard) for the alcohol+virus aetiology contrast. There were only 17 patients in the alcohol +virus category.

A related issue to the consequence of the small numbers of patients available is that of having a high degree of censoring. High proportions of censored data were present in the majority of the survival analysis studies undertaken in the present investigation, and this was particularly evident in the PBC study. The high degree of censoring was reflected in the instability of the model coefficients (as seen using the bootstrapping techniques). The **effective sample size** in a survival analysis model is often taken to be $N\psi$ where N is the sample size and ψ represents the proportion of uncensored

values (as in Schmoor et al, 2000). As mentioned above in the discussion of the PBC data, it is recommended that the EPV in a PH model are at least ten, with lower values resulting in an unreliable model. Most published studies do not meet this criterion, however (Altman & Royston, 2000)!

2) For most of the statistical analyses undertaken in the present thesis, as expected using statistical methods to derive prognostic models, the analyses were data-dependent rather than prespecified. It is known that data-driven methods are expected to provide an overoptimistic assessment of predictive performance. This problem of overoptimistic prediction has only recently come to light. In addition, the Cox models fitted to the data of the cirrhotic patients were based on sparse prognostic information and no validation data sets were available. The estimated Brier score and measures of residual variation employed may result in over-optimism when calculated in the same data from which the prognostic classification system has been derived (Graf et al, 1999). Computer intensive statistical techniques such as **bootstrapping** and **leave-one-out cross-validation,** attempt to reduce overoptimism at the model-building stage. They can also be used to estimate **shrinkage factors**, which can be applied to regression coefficients to counterbalance overoptimism (Altman & Royston, 2000, Schumacher et al, 1997). This new area will be examined in future research. Bootstrapping techniques can be used to investigate not only the stability of the variables included in a Cox model but also the estimated survival probabilities for individual patients (Altman & Andersen, 1989). This was not investigated in the present study due to the sparseness of prognostic information, but it is hoped that it will be investigated further using new data in the future.

Further approaches to model-fitting

The survival analyses undertaken in the present study have relied heavily on the Cox PH model, in which the hazard function in patient groups is compared to the baseline population using a multiplicative model on the log hazard scale. There is a second class of models that has been considered in the statistical literature, using which the survival functions are modelled directly on the time scale, thus accelerating or decelerating the time to failure. These are called accelerated failure time models (pg 14, Everitt & Dunn, 1998). Other than the PH model, the most popular model associating h(t) and x is the accelerated failure time model (pg31, DuCroq, 2000).

An alternative to the use of time-dependent models is to create multistate models. Altman & de Stavola (1994) state that it may be appropriate to consider applying a multistate model rather than a time-dependent model in cases in which the variables under consideration are few and discrete or identify complications which prelude death. The decompensation of cirrhosis patients could be taken to be such a variable. This type of model could also be applied in the case of PBC data in which transplantation could be taken into account as an intermediate event (as mentioned by Bonnand & Poupon, 1996).

The theory underlying the survival analyses has been described in a traditional manner in the present thesis, so as to be comprehendible by a wide range of health science professionals. Venables & Ripley (1994, pg267) mention that the modern mathematical approach to survival analysis is based on continuous parameter martingales. The general approach taken in the present study in obtaining prognostic models is the "classical approach", in that maximum likelihood methods have been applied in drawing inferences from the models fitted. An alternative that could be explored is that of a Bayesian approach, although there is general controversy surrounding the use of pre-specified prior distributions. Also, there are alternatives to the use of traditional survival analysis modelling methodologies, such as neural networks and regression trees (CART). These have only recently been explored in the literature and although there does not appear to be evidence that they offer any consistent advantage in the context of survival analysis (Altman & Royston, 2000), one personal aim is to explore this area further in the future.

It is hoped that the investigations undertaken in the present thesis provide a new insight into prognosis for Cretan cirrhosis patients. As the aim was to obtain clinical predictions for this previously unexamined group of subjects, rather than to delve into the statistical intricacies of various approaches to modelling, the models formed were based in the main on widely applicable traditional statistical methodology. Certain weaknesses in terms of the validity of specific models have been highlighted using computer-intensive techniques. The overall picture obtained suggests that there are interesting findings in terms of both therapeutic options and possibly useful clinical tools for this diverse group of patients.

6. REFERENCES

Agresti A (1984) *Analysis of Ordinal Categorical Data.* Wiley, New York.

Alexandrakis M, Coulocheri S, Kyriakou D, et al. (1997) Diagnostic value of ferritin, haptoglobin, a_1-antitripsin lactate dehydrogenase and complement factors C3 and C4 in pleural effusion differentiation. *Respir Med* **91**: 517-23

Alexandrakis MG, Coulocheri SA, Bouros D, Vlachonikolis IG & Eliopoulos GD (2000) Significance of alpha-2-macrogloboulin, alpha-1-acid glycoprotein, and c-reactive protein in pleural effusion differentiation. *Respiration* **67**:30-35

Alexandrakis MG, Moschandrea J, Kyriakou DS, Alexandraki R & Kouroumalis E. (2001) Use of a variety of biological parameters in distinguishing cirrhotic from malignant ascites. *Int J of Biol Markers* **16**(1): 45-9

Allgaier HP, Becker G & Blum HE (2000)A therapeutic study of hepatocellular carcinoma using octreotide (HECTOR) *Dtsch. Med. Wochenschr*. **125** (11):320

Altman DG (1998) Confidence intervals for the number needed to treat. *BMJ* **317**: 1309-12

Altman DG (2000) Statistics in medical journals: some recent trends. *Stat Med* **19**(23): 3275-89

Altman DG & Andersen PK (1989) Bootstrap investigation of the stability of a Cox regression model *Stat Med* **8**(7): 771-83

Altman DG & Andersen PK (1999) Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ* **319**(7223):1492-5

Altman DG & De Stavola BL (1994) Practical Problems in fitting a proportional hazards model to data with updated measurements of the covariates. *Stat Med* **13**: 301-41

Altman DG, De Stavola BL, Love SB & Stepniewska KA (1995) Review of survival analyses published in cancer journals. *Br J Cancer* **72(**2): 511-8

Altman DG & Royston P (2000) What do we mean by validating a prognostic model? *Stat Med* **19**(4): 453-73

Anthony PP, Ishak KG, Nayak NC et al. (1977) The morphology of cirrhosis: definition, nomenclature and classification. *Bull WHO* **55**: 521-540

Anthony PP, Ishak KG, Nayak NC et al. (1978) The morphology of cirrhosis. *J Clin Pathol* **31**: 395-414

Armitage P & Berry G (1987) *Statistical Methods in Medical Research* 2nd ed. Blackwell Oxford

Badvie S (2000) Hepatocellular carcinoma. *Postgrad Med J*. **76**(891): 4-11

Banke O, Dybkjaer E, Nordfelt E & Reineicke V (1971) Australia antigen and antibody in 10000 Danish blood donors. *Lancet* **1**: 860-61

Bansal S, Kaur K & Bansal AK (1998) Diagnosing ascitic biology on a biochemical basis. *Hepato-Gastroenterol* **45**: 1673-1677

Barlow WE & Prentice R (1988) Residuals for relative risk regression *Biometrika* **75**: 65-74

Benvegnu L, Fattovich G, Noventa F, Tremolada F, Chemello L, Cecchetto A & Alberti A (1994) Concurrent hepatitis B and C virus infection and risk of hepatocellular carcinoma in cirrhosis. A prospective study. *Cancer* **74**(9):2442-8.

Bergsland EK & Venook AP (2000) Hepatocellular carcinoma *Curr Opin Oncol* **12**:(4):357-61

Berk PD (2000) Hepatitis C virus (HCV). *Semin Liver Dis* **20**(1):iii-v

Bland M (2000) *An Introduction to Medical Statistics 3rd ed.* Oxford: Oxford University Press.

Bolondi L, Sofia S, Siringo S, Gaiani S, Casali A, Zironi G, Piscaglia F, Gramantieri L, Zanetti M & Sherman M (2001) Surveillance programme of cirrhotic patients for early diagnosis and treatment of hepatocellular carcinoma: a cost effectiveness analysis. *Gut* **48**(2): 251-9

Bonnand A-M, Heathcote EJ, Lindor KD & Poupon RE (1999) Clinical significance of serum bilirium levels under ursodeoxycholic acid therapy in patients with primary biliary cirrhosis. *Hepatology* **29**(1): 39-43

Bonnand AM & Poupon RE (1996) Survival in ursodeoxycholic acid-treated patients with biliary cirrhosis. *Gastroenterology* **111**(6) :1782

Bonis PAL, Tong MJ, Blatt LM, Conrad A & Griffith JL (1999) A predictive model for the development of hepatocellular carcinoma, liver failure, or liver transplantation for patients presenting to clinic with chronic hepatitis C. *Am J Gastroenterol* **94**(6):1605-12

Breslow NE (1974) Covariance analysis of censored data. *Biometrika* **2**: 437-53

Breslow NE & Day NE (1980) *Statistical Methods in Cancer Research* Volume 1. The Analysis of Case-Control Studies. Lyon: International agency for Research in Cancer (IARC) scientific publications No. 32

Bruix J, Barrera JM, Calvet X, Ercilla G, Costa J, Sanchez-Tapias JM, Ventura M, Vall M, Bruguera M, Bru C et al. (1989) Prevalence of antibodies to hepatitis C virus in spanish patients with hepatocellular carcinoma and hepatic cirrhosis. *Lancet* **28**(2):1004-6

Chambers JM & Hastie TJ (eds) (1993). *Statistical Models in S* Chapman & Hall, London

Chiba T, Matsuzaki Y, Abei M, et al. (1996) The role of previous hepatitis B virus infection and heavy smoking in hepatitis C virus related hepatocellular carcinoma. *Am J Gastroenterol* **6**:1195-1203

Chiaramonte M, Stroffolini D, Vian A, Stazi MA, Floreani A, Lorenzoni U, Lobello S, Farinati F & Naccarato R (1999) Rate of incidence of hepatocellular carcinoma in patients with compensated viral cirrhosis. *Cancer* **85 (**10**)**: 2132-7

Christensen E, Altman DJ, Neuberger J, De Savola,BL, Tygstrup N, Williams R (1993) Updating prognosis in primary biliary cirrhosis using a time-dependent Cox regression model. PBC1 and PBC2 trial groups. *Gastroenterology* **105**(6): 1865-76

Christensen E (1997) Prognostic models in chronic liver disease: validity, usefulness and future role. *J Hepatol* **26**(6): 1414-24

Collett D (1994) *Modelling Survival Data in Medical Research* Chapman & Hall, London

Colombo M, De Francis R, Del Ninno E, et al (1991) Hepatocellular carcinoma in Italian patients without cirrhosis. *N Engl J Med*.**325**: 675-680.

Colombo M, Choo QL, Donato MF, et al. (1989) Prevalence of antibodies to hepatitis C virus in Italian patients with hepatocellular carcinoma *Lancet* **ii**: 1006-8

Corrao G, Arico S (1998) Independent and combined action of hepatitis C virus infection and alcohol consumption on the risk of symptomatic liver cirrhosis: *Hepatology:* **27**(4): 914-9

Cox DR (1975) Partial likelihood. *Biometrika* **62**: 269-76

Cox DR (1972) Regression models and life tables. *J R Stat Soc* **34**: 248-75

Cox DR & Oakes D (1984) *Analysis of survival data.* Chapman & Hall, London

Cox DR & Snell EJ (1968) A generalised definition of residuals (with discussion). *Journal of the Royal Statistical Society, A,* **30**: 248-75

D' Amico G, Morabito A, Pagliaro L, Marubini E (1986) Survival and prognostic indicators in compensated and decompensated cirrhosis. *Dig Dis Sci.* **31** (5):468-75.

De Jongh FE, Janssen HJ, de Man RA, Hop WC, Schalm SW, van Blankenstein M, (1992). Survival and prognostic indicators in hepatitis B surface antigen-positive cirrhosis of the liver. *Gastroenterology* **103**(5): 1630-5

Dekker A, Bupp PA (1978) Cytology of serous effusions: an investigation into the usefulness of cell blocks versus smears. *Am J Clin Pathol* **70**:855-860

De Luca M, Ascione A, Vacca C, Zarone A (1992) Are healthcare workers really at risk of HCV infection? Letter *Lancet* **339**: 1364-5

Diamantis ID, McGandy CE, Pult I et al. (1992) Polymerase chain reaction detects hepatitis B virus DNA in paraffin-embedded liver tissue from patients sero-and histo-negative for active hepatitis B *Virchows Arch A Pathol Anat* **420:** 14-15

Di Bisceglie AM (1998) Hepatitis C (Seminar), *Lancet* **351**(9099): 351-55

Dickson ER, Grambsch PM, Fleming TR, Fisher LD, Langworthy A (1989) Prognosis in primary biliary cirrhosis: model for decision making. *Hepatology* **10**(1): 1-7

Dienstag JL (1997) The natural history of chronic hepatitis C and what we should do about it. *Amer. Gastroenterological Association* **112**(2): 651-55

Di Marco V, Lo Iacono O, Camma C, Vaccaro A, Giunta M, Martorana G, Fuschi P, Almasio PL, Craxi A (1999) The long-term course of chronic hepatitis B. *Hepatology* **30**(1):257-64

Ducrocq V. (2000) *Survival Analysis Applied to Animal Breeding & Epidemiology* Station de Genetique Quantitative et Appliquee, Institut National de la Recherche Agronomique France

Efron B (1977) The efficiency of Cox's likelihood function for censored data. *J Am Stat Assoc* **72**:557-65

Efron B (1982) *The Jacknife, Bootstrap, and Other Resampling Plans.* SIAM, Philadelphia.

Efron B & Tibshirani R. (1986) Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statist. Sci.* **1**:54-77

El-RefaieA, Savage K, Bhattacharva S Khakoo S, Harrison TJ, el-Batanony M, Soliman el-S, Nasr S, Mokhtar N, Amer K, Scheuer PJ, Dhillon AP (1996) HCV-associated hepatocellular carcinoma without cirrhosis. *J Hepatol* **24**:277-85

Everitt BS & Dunn G. (1998) *Statistical Analysis of Medical Data. New Developments* Arnold, London.

Everitt BS & Dunn G. (1991) *Applied Multivariate Data Analysis* Arnold, London.

Everhart JE & Hoofnagle JH (1992) Editorial comment on Hepititis B-related end-stage liver disease *Gastroenterology* **103**: 1692-4

Fasani P, Sangiovanni A, De Fazio C, Borzio M, Bruno S, Ronchi G, Del Ninno E & Colombo M (1999) High prevalence of multinodular hepatocellular carcinoma in patients with cirrhosis attributable to multiple risk factors. *Hepatology.* 29(6):1704-7.

Fattovich G (1998). Progression of hepatitis B and C to hepatocellular carcinoma in Western countries. *Hepatogastroenterology* **45** (Suppl 3):1206-13

Fattovich G, Giustina G, Schalm SW, Hadziyannis S, Sanchez-Tapias J, Almasio P, Christensen E, Krogsgaard K, Degos F, Carneiro de Moura M, et al (1995). Occurrence of hepatocellular carcinoma and decompensation in western European patients with cirrhosis type B. Cirrhosis. *Hepatology* **21**(1): 77-82

Fattovich G, Guistina G, Degos F, et al (1997) Morbidity and mortality in compensated cirrhosis type C: A retrospective follow-up stady of 384 patients. *Gastroenterology* **112**: 463-72

Fattovich G, Giustina G, Sanchez-Tapias J, Quero C, Mas A, Olivotto PG, Solinas A, Almasio P, Hadziyannis S, Degos F, de Moura MC, Krogsgaard K, Pantalena M, Realdi G, Corrocher R, Schalm SW (1998) Delayed clearance of serum HBsAg in compensated cirrhosis B: relation to interferon alpha therapy and disease prognosis. European Concerted Action on Viral Hepatitis (EUROHEP) *Am J Gastroenterol.* **93**(6):896-900.

Fattovich G, Giustina G, Christensen E, Pantalena M, Zagni I, Realdi G, Schalm SW, (2000). Influence of hepatitis delta virus infection on morbidity and mortality in compensated cirrhosis type B. The European Concerted Action on Viral Hepatitis. *Gut* **46**(3): 420-6

Fattovich G, Ribero ML, Pantalena M, Diodati G, Almasio P, Nevens F et al., Eurohep Study Group on Viral Hepatitis (2001), Hepatitis C virus genotypes: distribution and clinical significance in patients with cirrhosis type C seen at tertiary referral centres in Europe. *J Viral Hepat* **8**: 206-16

Fattovich G, Brollo L, Giustina G, Noventa F, Pontisso P, Alberti A, Realdi G, Ruol A (1991) Natural history and prognostic factors for chronic hepatitis type B. **32**(3): 294-8.

Fernandez-Esparrach G, Sanchez-Fueyo A, Gines P, Uriz J, Quinto L, Ventura PJ, Cardenas A, Guevara M, Sort P, Jimenez W, Bataller R, Arroyo V, Rodes J (2001) A prognostic model for predicting survival in cirrhosis with ascites. *J Hepatol* **34**(1): 46-52

Fragiadakis N, Carahristos A, Manoussos O et al. (1996) Prevalence of anti-HCV and anti-HEV in the district of Heraklion of Crete-Greece: an urban and rural population based study. *IX International Symposium on Viral Hepatitis and Liver Disease*. Rome Abstracts: 188

Galambos JT (1985) *Epidemiology of alcoholic liver disease in the United States of America.* In Hall P de la M. ed. *Alcoholic liver disease.* Eduward Arnold. London

Garrison RN, Kaelin LD, Hauser LS, Galloway RH (1986) Malignant ascites: Clinical and experimental observations. *Ann Surg* **203**:644-651

Gerbes AL, Jungst D, Xie Y, Permanetter W, Paumgartner G (1991**)** Ascitic fluid analysis for the differentiation of malignancy - related and nonmalignant ascites *Cancer* 68:1808-1814

Gines P, Quintero E, Arroyo V, et al (1987) Compensated cirrhosis: Natural history and prognostic factors. *Hepatology* **7**:122-8

Gonzalez A, Esteban JI, Madoz P, Viladomiu L, Genesca J, Muniz E, Enriquez J, Torras X, Hernandez JM, Quer J, et al. (1995) Efficacy of screening donors for antibodies to hepatitis C virus to prevent transfusion-assosiated hepatitis: final report of a prospective trial. *Hepatology* **22**: 439-45

Graf E, Schmoor C, Sauerbrei W, Schumacher M (1999) Assessment and comparison of prognostic classification schemes for survival data. *Stat Med* **18**(17-18): 2529-45

Grambsch PM, Dickson ER, Wiesner RH, Langworthy A (1989) Application of the Mayo primary biliary cirrhosis survival model to Mayo liver transplant patients. *Mayo Clin Proc* **64**(6): 699-704

Greenman Y & Melmed S (1994) Expression of three somatostatin receptor subtypes in primary adenomas: evidence for preferential SSTR5 expression in the mammosomatotroph lineage. *J Clin Endocrinol Metab* **79**:724-9

Greenwood M (1926) The errors of sampling of the survivorship tables, in *Reports on Public Health and Statistical subjects* **33**:Appendix 1, HMSO, London

Goulis J, Leandro G, Burroughs AK (1999) Randomised controlled trials of ursodeoxycholic acid therapy for primary biliary cirrhosis: a metam analysis. *Lancet* **354** : 1053-60

Gupta R, Misra SP, Dwivedi M, Misra V, Kumar S, Gupta SC (1995) Diagnosing ascites: Value of ascitic fluid total protein, albumin, cholesterol their ratios, serum-ascites albumin and cholesterol gradient. *J. Gastroenterol Hepatol* **10**:295-299

Hadziyiannis SJ, (1980) Hepatocellular carcinoma and type B hepatitis. *Clin Gastrenterol* **9**: 117-34

Hansson BG, Sundstrom G, Johnsson T, (1975) Occurrence of hepatitis B surface antigen (HBsAg) in blood donors and patients suffering from transfusion hepatitis. *Acta Pathol Microbiol Immunol Scand Sect C* **83**: 112-20

Hatzakis A, Katsoulidou A, Kaklamani E, et al. (1996) Hepititis C virus 1b is the dominant genotype in HCV-related carcinogenesis: A case-control study. *Int J Cancer* **68**:51-3

Hayes PC & Simpson KJ (1995) *Gastroenterology and Liver Disease* Churchill Livingstone, Edinburgh

Heathcote EJ (2000) Management of primary biliary cirrhosis. *Hepatology* **31**:1005-13.

Heathcote J (2001) Primary Biliary Cirrhosis. *Clinical Perspectives in Gastroenterology* January/February 39-46

Heathcote EJ, Cauch-Durek K, Walker V, Bailey RJ, Blendis LM et al.(1994) The Canadian multicenter double-blind randomized controlled trial of ursodeoxycholic acid in primary biliary cirrhosis. *Hepatology* **19**(5): 1149-56

Hinkley DV, Reid N, Snell EJ eds (1991), S*tatistical Theory and Modelling: In honour of Sir David Cox FRS* Chapman and Hall, London

Hislop WS, Masterton N, Bouchier IAD, Hopwood D.(1982) Cirrhosis and primary liver cell carcinoma in Tayside: a five year study. *Scot Med J* **27**: 29-36

Ikeda K, Saitoh S, Koida I, Arase Y, Tsubota A, Chayama K, Kumada H, Kawanishi M (1993) A multivariate analysis of risk factors for hepatocellular carcinogenesis: a prospective observation of 795 patients with viral and alcoholic cirrhosis *Hepatology* **18**(1):47-53

Iwarsson S, Lindholm A, Lundin P et al., (1972) Hepatitis associated antigen and antibody in Swedish blood donors. Relation to liver disease. *Vox Sang* **22**: 501-9

Jones DEJ, Metcalf JV, Collier JD, Bassendine MF, James OFW (1997) Hepatocellular carcinoma in primary biliary cirrhosis and its impact on outcomes. *Hepatology* **26**(5): 1138-42

Jungst D, Gerbes AL, Martin R, Paumgartner G (1986) Value of ascitic lipids in the differentiation between cirrhotic and malignant ascites. *Hepatology* **6**: 239-244

Kaplan GL & Meier P (1958) Non-parametric estimation from incomplete observations. *J Am Stat Assoc* **53**: 457-81

Kew MC, Houghton M Choo QL, et al. (1990) Hepatitis C virus antibodies in southern African blacks with hepatocellular carcinoma. *Lancet* **335**: 873-4

Kim WR, Wiesner RH, Poterucha JJ, Therneau TM, Benson JT, Krom RA, Dickson ER (2000) Adaptation of the Mayo primary biliary cirrhosis natural history model for application in liver transplant candidates. *Liver Transpl* **6**(4): 489-94

Kiyosawa K, Sodevama T, Tanaka E et al. (1990) Interrelationship of blood transfusion, non-A, non-B hepititis and hepatocellular carcinoma: analysis by detection of antibody to hepatitis C virus. *Hepatology* **12**:671-5

Koulentaki M, Ergazaki M, Moschandrea J, Spanoudakis S, Tzagarakis N, Drandakis PE, Spandidos DA, Kouroumalis EA. (2001) Prevalence of hepatitis B and C markers in high-risk hospitalised patients in Crete: a five-year observational study. *BMC Public Health*. 1(1):17.

Koulentaki M, Spanoudakis S, Kantidaki E, Drandakis P, Tzagarakis N, Biziakos, E Moschandrea J, Kouroumalis EA (1999) Prevalence of hepatitis B and C markers in volunteer blood donors in Crete: A 5-year study. *J Viral Hep* **6**:1-6

Kouroumalis EA, Skordilis PG, Moschandrea I, et al (1997) Natural history of advanced hepatocellular carcinoma in Crete. Association with hepatitis C virus. *Eur J Gastroent Hepatol* **9**(10): 981-8

Kouroumalis EA, Skordilis PG, Thermos K, Vasilaki A, Moschandrea J, Manousos ON (1998) Treatment of hepatocellular carsinoma with ocreotide: a randomised controlled study. *Gut* **42:** 442-47

Krzeski P, Zych W, Kraszewska E, Milewski B, Butruk E, Habior (1999) Is serum bilirubin concentration the only valid prognostic marker in primary biliary cirrhosis? *Hepatology* **30**(4): 865-9

Kyriakis KP, Foudoulaki IE, Papoulia EI, Sofroniadou KE. (2000) Seroprevalence of hepatitis B surface antigen (HbsAg) among first-time and sporadic blood donors in Greece: 1991-1996 *Transfusion Medicine* **10**:175-80

Laupacis A, Sackett DL, Roberts RS (1988) An assesment of clinically useful measures of the consequences of treatment. *New Eng J Med* **318**: 1728-33

Lee CM, Changhien CS, Shyu WC, Liaw YF (1992) Serum ascites albumin concentration gradient and ascites fibronectin in the diagnosis of malignant ascites. *Cancer* **70**: 2057-2060

Lee ET (1992) *Statistical Methods for Survival Analysis* John Wiley & Sons. New York.

Lee FI (1966) Cirrhosis and hepatoma in alcoholics. *Gut* **7**: 77-85

Lee H-S, Hans CI, Kim CY (1993) Predominant etiologic association of hepatitis C virus with hepatocellular carcinoma compared with hepatitis B virus in elderly patients in a hepatitis B-endemic area. *Cancer* **72:** 2564-7

Lindor KD, Dickson ER, Baldus WP, Jorgensen RA, Ludwig J, Murtaugh P, Harrison JM, Wiesner RH, Anderson ML, Lange SM, Lesage G, Rossi SS, Hofmann AF (1994) Ursodeoxycholic acid in the treatment of primary biliary cirrhosis. *Gastroenterology* **106**: 1284-90

Lindor KD, Therneau TM, Jorgensen RA, Malinchoc M, Dickson ER (1996) Effects of ursodeoxycholic acid on survival in patients with primary biliary cirrhosis. *Gastroenterology* **110**(5): 1515-8

Lindor KD & Therneau TM (1996) Reply to letter: Survival in ursodeoxycholic acid-treated patients with biliary cirrhosis by Bonnand AM & Poupon RE. *Gastroenterology* **111**(6): 1782-3

Lionis C, Frangoulis E, Koulentakis M, Biziagos E, Kouroumalis E (1997a). Prevalence of hepatitis A, B, and C markers in school children of a rural area of Crete, Greece. *Eur J Epidemiol* **13**(4): 417-20

Lionis C, Koulentaki M, Biziagos E, et al. (1997b) Current prevalence of hepatitis A, B, and C markers in a well defined area in rural Crete, Greece. *J Viral Hepat* **4** :55-61

Lionis C, Vlachonikolis IG, Skliros S, Symeonidis A, Merkouris BP, Kouroumalis E.(2000). Do undefined sources of hepatitis C transmission exist? The Greek study in General Practice. *J Viral Hepat* **7**(3):218-24

Liovet JM, Sala M, Castells L, et al. (2000) Randomised controlled trial of IFN treatment for advanced HCC. *Hepatology* **31**(1): 54-8

Little RJ & Rubin DB (1987) *Statistical Analysis with Missing Data* John Wiley & Sons Inc, New York

Liu CL, Fan ST, Ng FO, et al (2000) Treatment of advanced HCC with tamoxifen and the correlation with expression of hormone receptors: a prospective randomised study. *Am. J. Gastroenterol.* **95**(1): 218-22

Ludwig J, Dickson ER, McDonald GSA (1978) Staging of chronic non-supperative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virch Arch A Pathol anat Histopathol* **379**:103-22

Machin D & Gardner MJ (1989) Calculating confidence intervals for survival time analysis, in *Statistics with Confidence-confidence intervals & statistical guidelines.* Garder MJ & Altman DG London: BMJ pp64-70

MacSweeny RNM, Anthony PP, Scheuer PJ, Burt AD, Portmann BC (eds) (1994) *Pathology of the Liver* Churchill Livingstone, Edinburgh.

Manesis EK, Giannoulis G, Zoumboulis P, Vafiadou I, Hadziyannis SJ (1995). Treatment of hepatocellular carcinoma with combined suppression and inhibition of sex hormones: a randomized, controlled trial. *Hepatology* **21**(6): 1535-42

Mantel N & Haenszel W (1959) Statistical aspects of the analysis of datafrom retrospective studies of disease. *J Nat Cancer Inst* **22**:719-48

Markus BH, Dickson ER, Grambsch PM,Fleming TR, Mazzaferro V, et al (1989) Efficacy of liver transplantation in patients with primary biliary cirrhosis. *N Eng J Med* **320**:1709-13

MathSoft Inc. (1997) *S-Plus 4: Guide to Statistics*. MathSoft Inc Seattle, Washington.

Mauer K, Manzione NC (1988) Usefulness of serum -ascites albumin difference in separating transudative from exudative ascites. Another look. *Dig Dis Sci* **33**:1208-1212

Meliconi R, Andreone P, Fasano L, Galli S, Pacilli A, Miniero R, Fabbri M,

Solforosi L & Bernardi M (1996) Incidence of hepatitis C virus infection in Italian patients with idiopathic pulmonary fibrosis. *Thorax* **51**(3):315-7.

Michel S, Sue B, Montpeyroux F, et al. (1997) Liver resection or transplantation for hepatocellular carcinoma? Retrospective analysis of 215 patients with cirrhosis *J Hepatol* 1274-80.

Moser C & Kalton G (1979) *Survey Methods in Social Investigation 2nd ed.* Gower. Great Britain

Mounoz-Gomez R, Garcia Mouzou CO, Garcia-Buey L, Lo Iacono O, Borque MJ, Garia-Sanchez A, Pajares JM, Moreno-Otero R (1996) Hepatitis C virus infection in Spanish volunteer blood donors: HCV RNA analysis and liver disease. *Eur J Gastroenterol Hepatol* **8**: 273-77

Neuberger J, Altman DG, Christensen E, Tygestrup N, Williams R (1986) Use of a prognostic index in evaluation of liver transplantation for primary biliary cirrhosis. *Transplantation* **41**(6): 713-6

Norušis MJ/SPSS Inc (1994) *SPSS Professional Statistics 6.1* SPSS Inc. USA

Nousbaum JB, Pol S, Naplas B, et al (1995) Hepatitis C virus type 1b (II) infection in France and Italy. *Ann Inter Med* **122** :161-8

Oakes, D (1991) Life table analysis, in S*tatistical Theory and Modelling: In honour of Sir David Cox FRS* (eds Hinkley DV, Reid N & Snell EJ), London: Chapman & Hall pp 107-28

Okuda H, Obata H, Motoike Y, Hisamitsu T. (1984) Clinicopathological features of hepatocellular carcinoma--comparison of hepatitis B seropositive and seronegative patients. *Hepatogastroenterology* **31**(2):64-8.

Parés A, Caballeria L, Rodes J, Bruguera M, Rodrigo L, Garcia-Plaza A, Berenguer J, Rodriguez-Martinez D, Mercader J, Velicia R (2000) Long-term effects of ursodeoxycholic acid in primary biliary cirrhosis: results of a double-blind controlled multicentric trial. UDCA-Cooperative Group from the Spanish Association for the Study of the Liver. *J Hepatol* **32**(4):561-6.

Parota M, Deloro K,Giordano M et al. (1996) Prevalence of anti-HCV in blood donors. *Ital J Gastroenterol* **23**: 555-8

Paterini P, Driss F, naplas B, et al. (1993) Persistence of hepatitis C viral genomes in primary liver cancer from HbsAg-negative patients: a study of a low endemic area. *Hepatology* **17**: 20-29

Peduzzi P, Concato J, Feinstein AR, Holford TR (1995) The importance of events per independent variable (EPV) in proportional hazards regression analysis II. Accuracy and precision of regression estimates. *J Clin Epidemiol* **48**: 1503-10

Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* **49**:1373-9.

Peto J (1984) The calculation and interpretation of survival curves, in: *Cancer clinical trials: methods and practice* (eds Buyse ME, Staquel MJ & Sylvester RJ) Oxford University Press pp361-80

Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG (1977) Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples *Br J Cancer* **35**(1):1-39

Philosophe B, Greig PD, Hemming AW, et al (1998) Surgical management of hepatocellular carcinoma: resection or transplantation? *J Gastrointest Surg* **2**: 21-7

Poupon RE, Balkau B, Gueshot J, Heinzmann F (1994) Predictive factors in ursodeoxycholic acid-treated patients: role of serum markers of connective tissue. *Hepatology* **19**: 635-40

Poupon RE, Bonnand AM, Chretien Y, Poupon R (1999) Ten-year survival in ursodeoxycholic acid-treated patients with primary biliary cirrhosis. *Hepatology* **29**(6): 1668-71

Poupon RE, Lindor KD, Cauch-Dudec K, Dickson ER, Poupon R, Heathcote EJ (1997) Combined analysis of randomized controlled trials of ursodeoxycholic acid in primary biliary cirrhosis. *Gastroenterology* **113**(3): 884-90

Poupon RE, Poupon R, Balkau B and the UDCA-PBC study group, (1994) Ursodiol for the long-term treatment of primary biliary cirrhosis. *N Engl J Med* **330**: 1342-47

Prentice RL & Marek P (1979) A quantitative discrepancy between censored data rank tests. *Biometrics* **35**:861-7

Quenouille MH (1956) Notes on bias in estimation *Biometrika* **43**:353-60

Raderer M, Hejna MH, Kurtaran A,et al (1999) Successful treatment of an advanced hepatocellular carcinoma with long acting somatostatin analogue lanreotide. *Am J Gastroenterol* **94**(1): 278-9

Raoul J.L, Boucher E, Kebrat P (1999) Non surgical treatment of hepatocellular carcinoma *Bull. Cancer,* **86**(6): 537-43

Realdi G, Fattovich G, Hadziyannis S, Schalm SW, Almasio P, Sanchez-Tapias J, Christensen E, Giustina G, Noventa F (1994) Survival and prognostic factors in 366 patients with compensated cirrhosis type B: a multicenter study. *J Hepatol* **21**(4): 656-66

Reubi JC, Kvols L, Krenning E, Lamberts SW (1990). Distribution of somatostatin receptors in normal and tumor tissue. *Metabolism* **39**(9 Suppl 2): 78-81

Reubi JC, Schaer JC, Waser B, Mengod G (1994) Expression of localization of somatostatin receptor SSTR1, SSTR2 and SSTR3 messenger RNAsin primary human tumours using in situ hybridization. *Cancer Res* **54:** 3455-9

Rice JA (1988) *Mathematical Statistics and Data Analysis* Wadsworth & Brooks Cole California.

Roudot-Thoraval F, Bastie A, Pawlotsky JM, et al (1997) Epidemiological factors affecting the severity of hepitits C virus-related liver disease: A French survey of 6,664 patients. The study group for the Prevalence and the Epidemiology of Hepititis C Virus. *Hematology* **26**: 485-90

Salmeron FJ, Palacios A, Perez-Ruiz M Torres C, Oyonarte S, Fernandez-Montoya A & Ruiz-Extremera A (1996) Epidemiology, serological markers and hepatitic disease of anti-HCV ELISA 2 positive blood donors. *Dig Dis Sci* **41**: 1933-98

Schally AV (1988) Oncological applications of somatostatin analogues. *Cancer Res* **48**: 6977-85

Scheuer FJ (1967) Primary biliary cirrhosis. *Proc R Soc Med* **60**: 1257-60

Schmoor C, Sauerbrei W, Schumacher M (2000) Sample size considerations for the evaluation of prognostic factors in survival analysis. *Stat Med* **19**(4): 441-52

Schneider D, Halperin R, Langer R, et al. (1997) Peritoneal fluid lactate dehydrogenase in ovarian cancer. *Gynecol Oncol* **66**: 399-94

Schoenfeld DA (1981) The asymptotic properties of nonparametric tests for comparing survival distributions. *Biometrika* **68**: 316-19

Schoenfeld DA (1982) Partial residuals for the proportional hazards regression model. *Biometrika* **69**:239-41

Scholmerich J, Volk BA, Kottgen E. Ehlerls S, Gerok W **(**1984) Fibronectin concentration in ascites differentiates between malignant and non malignant ascites. *Gastroenterology* **87**:1160-1164

Schumacher M, Hollander N, Sauerbrei W (1997). Resampling and cross-validation techniques: a tool to reduce bias caused by model building? *Stat Med* **16**(24): 2813-27

Segal MR (1998) Tree-structured survival analysis in medical research, in *Statistical Analysis of Medical Data* (eds Everitt BS & Dunn G), Arnold, London pp101- 125

Serfaty L, Aumaitre H, Chazouilleres O, Bonnand AM, Rosmotruc O, Poupon RE & Poupon R (1998) Determinants of outcome of compensated hepatitis C virus-related cirrhosis. Comment in *Hepatology* **27**(5): 1435-40

Shapiro JM, Smith H & Schaffner F (1979) Serum bilirubin: a prognostic factor in primary biliary cirrhosis. *Gut* **20**:137-40

Shijo H, Okazaki M, Koganemaru F, Higashi M, Sakaguchi S, Okumura M (1991) Influence of hepatitis B virus infection and age on mode of growth of hepatocellular carcinoma. *Cancer* 15 (67):2626-32.

Shouval D (1998) Octeotide in hepatocellular carcinoma, *Gut* **42**(3): 136-7

Shuster JJ (1991) Median follow up in clinical trials. *J Clin Oncol,* **9:**191-192

Sokal RR, Rohlf FJ (1995) *Biometry* 3rd ed. Freeman WH & Co, New York

Sφrensen TIA, Bensen KD, Eghoje KN (1984) Prospective evaluation of alcohol abuse and alcoholic liver injury in men as predictors of development of cirrhosis. *Lancet* **2** (8397): 241-44

Simoneti RG, Liberati A, Angiolini C, et al (1997) Treatment of HCC: A systematic review of randomised controlled trials. *Ann Oncol* **8**(2): 117-36.

SPSS Inc. (1997) *SPSS Advanced Statistics 7.5* SPSS Inc. USA

Stablein DM, Walter H, Carter Jr, Novak JW (1981) Analysis of survival data with nonproportional hazard functions. *Controlled Clin Trials* **2**: 129-59

Tanaka K, Ikematsu H, Hirchata T, et al. (1996) Hepitits C virus infection and risk of hepatocellular carcinoma among Japanese: possible role of type1b(II) infection. *J Natl Cancer Inst* **68**:742-46

Therneau TM, Grambsch PM & Fleming TR (1990) Martingale-based residuals for survival models. *Biometrika* **77**: 147-60

Thompsen BL, Keiding N & Altman DG (1991) A note on the calculation of expected survival, illustrated by the survival of liver transplant patients. *Stat Med* **10**:733-8

Tremolada F, Casarin C, alberti A, Drago C, Tagger A, Ribero ML, Realdi G. (1992) Long-term follow-up of non-A, non-B (type) C post-transfusion hepatitis. *J Hepatol* **16**: 273-81

Trichopoulos D, Vidaki M, Sparros L, Hiraikaki E (1975) Epidemiology and hepatitis B and primary hepatic carcinoma (P.H.C.) *Lancet* **2**(2943): 1038-9

Trichopoulos D, Talbor E, Gerety RJ, et al. (1978) Hepatitis B and primary hepatocellular carcinoma in a European population. *Lancet* **2**(8102): 1217-9

Tsukuma H, Hiyama T, Tanaka S, Nakao M, Yabuuchi T, Kitamura T, Nakanishi K, Fujimoto I, Inoue A, Yamazaki H et al. (1993) Risk factors for hepatocellular carcinoma among patients with chronic liver disease. *N Engl J Med* 24 (328):1797-801

Venables WN & Ripley BD (1994) *Modern Applied Statistics with S-Plus* Springer-Verlag NY, Inc USA

Villamil FG, Sorroche PB, Aziz HF, Lopez PM & Oyhamburu JM (1990) Ascitic fluid a_1-antitrypsin. *Dig Dis Sci* **35**: 1105-9

Vissoulis H, Hadziyannis S, Papaevagelou G, Afroudakis A, Giannopoulos C Gioustose A, Merikas G (1972) Geographical distribution of hepatitis associated antigen (HAA) carriers in Greece. *Nosokomiaka Chronika **34**:* 5-10 (in greek)

Voeth M (1992) Letter to the Editor, on: A note on the calculation of expected survival, illustrated by survival of liver transplant patients by Thomsen BL et al, *Stat Med* **11**: 1527-32

Wang SS, Lee FY, Chao Y et al. (1996) Clinical significance of complements in ascitic diseases: Elevated complement levels disapproving the liver disease origin. *Proc Natl Sci Counc Repub China B* **20**: 51-7

Wasley A & Alter MJ (2000) Epidemiology of hepatitis C: geographic differences and temporal trends. *Semin Liver Dis* **20**(1): 1-16

WHO (1999) Global surveillance and control of hepatitis C. Report of a WHO Consultation organised in collaboration with the Viral Hepatitis Prevention Board, *Antwerp*, Belgium. *J Viral Hepat* 6:35-47

WHO (2000) Hepatitis B: Fact Sheet No 204

WHO (2000) Hepititis C: Fact Sheet No 164

Wiesner RH (1998) Liver transplantation for primary biliary cirrhosis and primary sclerosing cholangitis: predicting outcomes with natural history models. *Mayo Clin Proc* **73**(6): 575-88

Woolson RF (1981) Rank tests and a one-sample log rank test for comparing observed survival data to a standard population. *Biometrics* **37**:632-640

Yamamoto S, Konishi I, Mandai M,Kuroda H, Komatsu T, Nambu K (1997) Expression of vascular endothelial grouth factor (VEGF) in epithelial ovarian neoplasms: correlation with clinicopathology and patient survival and analysis of serum VEGF levels. *Br J Cancer* **76:** 1221-1227.

Yoshida H, Hamada T, Inuzuka S, Ueno T, Sata M, Tanikawa K (1993) Bacterial infection in cirrhosis, with and without hepatocellular carcinoma. *Am J Gastroenterol* **88**(12): 2067-71

Zein NN, Poteruka JJ, Gross JB Jr et al (1996) Increased risk of hepatocellular carcinoma in patients infected with hepatitis C genotype 1b. *Am J Gastroenterol* **91**:2560-2

Zervou EK, Dalekos GN, Boumba DS, Tsianos EV (2001) Value of anti-HBc screening of blood donors for prevention of HBV infection: results of a 3-year

prospective study in Northwestern Greece. *Transfusion* 41(5): 652-8

7.1 APPENDIX A. Survival analysis in SPSS and S-PLUS

I) Cox regression in SPSS

When fitting a Cox regression model in SPSS, syntax of a similar format to the following may be used:

```
COXREG
  survtime  /STATUS=event(1)
 /CONTRAST (sex)=Indicator(1)
  /METHOD=BSTEP(LR) sex age albu bili
  /PRINT=CI(95) SUMMARY BASELINE
  /SAVE SURVIVAL HAZARD XBETA
  /PATTERN AGE(50)
  /CRITERIA=PIN(.05) POUT(.10) ITERATE(20)
  /OUTFILE=COEFF(mdlcffs) TABLE(mrfns).
```

The SAVE subcommand allows one to save the linear combination of mean corrected covariate values multiplied by regression coefficients from the final model (XBETA)

The printout includes an estimate of the baseline cumulative hazard at the observed time point (i.e. survival time) of each individual. This is the quantity $\hat{H}_0(t_i)$.

Also provided are estimates of the survivor function, the s.e. of the estimate and the estimated cumulative hazard at baseline and also when all variables are at their mean. The means used are also provided in the output. These variables can all be obtained and saved using the OUTFILE command.

II) An example of an S-Plus survival analysis session: HCC treated patients versus historical controls

Kaplan-Meier analysis and log rank test

```
> attach(Histcont20cntrls)
> Surv(TIME,1-EVENT)
 [1]  7.016393+ 22.983607  26.950820+ 22.983607  14.983607+ 23.967213+
24.983607   6.983607  18.000000+
[10]  4.000000+  8.000000+  3.934426  11.016393   2.000000  11.967213
5.016393   4.983607  13.967213
[19]  4.000000   4.032787  14.983607   5.934426  12.950820  14.983607+
2.983607+  9.016393+  5.016393+
[28]  8.000000+ 18.000000+ 32.983607+  4.000000  24.950820+  6.000000
4.000000  16.000000  16.000000
[37] 33.000000   6.000000   9.000000   5.000000   5.000000   3.000000
4.000000   3.000000   8.000000
[46]  4.000000  10.000000  25.000000+  5.000000  14.000000  22.000000
9.000000
> histcont20.surv<-survfit(Surv(TIME,1-EVENT)~PATTYPE,conf.type="log-log")
> summary(histcont20.surv)
Call: survfit(formula = Surv(TIME, 1 - EVENT) ~ PATTYPE, conf.type = "log-
log")

                PATTYPE=1
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
  2.00     32       1    0.969  0.0308       0.7982        0.996
  3.93     30       1    0.936  0.0435       0.7690        0.984
  4.00     29       2    0.872  0.0598       0.6937        0.950
  4.03     26       1    0.838  0.0663       0.6540        0.929
  4.98     25       1    0.805  0.0716       0.6158        0.907
  5.02     24       1    0.771  0.0761       0.5788        0.884
  5.93     22       1    0.736  0.0803       0.5404        0.859
  6.98     21       1    0.701  0.0838       0.5033        0.832
 11.02     16       1    0.657  0.0893       0.4533        0.801
 11.97     15       1    0.614  0.0935       0.4065        0.767
 12.95     14       1    0.570  0.0965       0.3624        0.732
 13.97     13       1    0.526  0.0985       0.3204        0.696
 14.98     12       1    0.482  0.0996       0.2805        0.658
 22.98      7       2    0.344  0.1088       0.1487        0.551
 24.98      3       1    0.230  0.1185       0.0536        0.477

                PATTYPE=2
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    3     20       2     0.90  0.0671       0.6560        0.974
    4     18       3     0.75  0.0968       0.4999        0.887
    5     15       3     0.60  0.1095       0.3573        0.776
    6     12       2     0.50  0.1118       0.2713        0.692
    8     10       1     0.45  0.1112       0.2311        0.647
    9      9       2     0.35  0.1067       0.1566        0.552
   10      7       1     0.30  0.1025       0.1225        0.501
   14      6       1     0.25  0.0968       0.0910        0.449
   16      5       2     0.15  0.0798       0.0373        0.335
   22      3       1     0.10  0.0671       0.0170        0.272
   33      1       1     0.00      NA           NA           NA

> plot(histcont20.surv, conf.int=T,lty=c(3,2),log=T,xlab="Survival
time(months)", ylab="Survival probability")
> legend(25,0.9,c("control","treated"),lty=c(2,3),lwd=2)
> survdiff(Surv(TIME,1-EVENT)~PATTYPE)
Call:
survdiff(formula = Surv(TIME, 1 - EVENT) ~ PATTYPE)

           N Observed Expected (O-E)^2/E (O-E)^2/V
PATTYPE=1 32       17     22.9      1.52      4.58
PATTYPE=2 20       19     13.1      2.65      4.58
```

```
 Chisq= 4.6  on 1 degrees of freedom, p= 0.0324
>

#Same as SPSS results, but here we have the expected no. too!
```

Cox PH modelling

```
> attach(Histcont20cntrls)
> plot(histcont20.surv,lty=c(3,4), xlab="Survival time
(months)",ylab="H(t)")
> histcont20.cox<-coxph(Surv(time,1-event)~pattype+okiandii)
#The Efron approximation is used as the default here (not the Breslow method
#as used in most applications), as it is much more accurate when
#dealing with tied death times, and is as efficient computationally.
> summary(histcont20.cox)
Call:
coxph(formula = Surv(time, 1 - event) ~ pattype + okiandii)
  n= 52
         coef exp(coef) se(coef)    z       p
 pattype 1.18      3.25    0.379 3.12 0.00180
okiandii 1.70      5.49    0.440 3.87 0.00011

         exp(coef) exp(-coef) lower .95 upper .95
 pattype      3.25      0.307      1.55      6.83
okiandii      5.49      0.182      2.32     13.01

Rsquare= 0.345   (max possible= 0.988 )
#The R2 measure is taken from Nagelkirke (1991) pg281
Likelihood ratio test= 22  on 2 df,   p=0.000017
Wald test            = 18.3  on 2 df,   p=0.000109
Efficient score test = 20.4  on 2 df,   p=0.0000367

> plot(survfit(histcont20.cox),lty=2:3,lwd=2,add=T,log=T)
#The above plot gives a single line + CIS
#Using strata separates the groups for the graph
#Separate baseline hazards are estimated for each stratum
> histcont20.coxs<-coxph(Surv(time,1-event)~strata(pattype)+okiandii)
> summary(histcont20.coxs)
Call:
coxph(formula = Surv(time, 1 - event) ~ strata(pattype) + okiandii)

  n= 52
         coef exp(coef) se(coef)    z      p
okiandii 1.84      6.28    0.508 3.62 0.0003

         exp(coef) exp(-coef) lower .95 upper .95
okiandii      6.28      0.159      2.32        17

Rsquare= 0.287   (max possible= 0.967 )
Likelihood ratio test= 17.6  on 1 df,   p=0.0000271
Wald test            = 13.1  on 1 df,   p=0.000297
Efficient score test = 16.3  on 1 df,   p=0.0000551

> plot(survfit(histcont20.coxs),lty=2:3,lwd=2,add=T,log=T)
> plot(survfit(histcont20.coxs),conf.int=T,lty=2:3,lwd=2,add=T,log=T)
#The CIs are very wide for the control group
```

7.2 APPENDIX B. Contrasts for Cox models using SPSS

In fitting the Cox models, sets of contrast variables are set up in order to compare levels of categorical variables such as sex and cirrhosis aetiology. The contrasts used in the present study are either indicator contrasts or simple contrasts. **Indicator contrasts** involve the setting up of dummy variables. If the categorical variable has k levels, k-1 dummy variables are set up. Cases in the reference category are coded 0 for all indictor variables except the ith, which is coded 1. When using **simple contrasts**, each category of the variable is compared to the reference category. For the purposes of Cox modelling, the same model coefficients are provided whether simple or indicator contrasts are used.

The SPSS procedure for estimation of the baseline cumulative hazard, however, prints a different baseline cumulative hazard (/PRINT=BASELINE) according to whether or not indicator contrasts have been used. Other types of contrasts (eg deviation constrasts) lead to the same baseline cumulative hazard as that produced using simple contrasts. This is a similar situation to logistic regression, in which changing the type of contrast used results only in a change of constant term in the model (the constant being the equivalent of the baseline hazard here) not in the odds ratios obtained. The baseline cumulative hazard produced using simple contrasts has been used in the present study for models involving categorical variables.

7.3 APPENDIX C. SPSS syntax for time dependent covariates

CLEAR TIME PROGRAM.

```
TIME PROGRAM.
COMPUTE rik = (T_>TmToRiks).
IF missing(TmToRiks) rik=0.
COMPUTE cancer = (T_>TmDghcc).
IF missing(TmDghcc) cancer=0.
COXREG
  SrvTmDg  /STATUS=died(1)
  /CONTRAST (typgpd2)=Indicator(1)/CONTRAST
(sex)=Indicator(1)
  /METHOD=BSTEP(LR) typgpd2 sex agediag rik cancer
 /PRINT=CI(95)
  /CRITERIA=PIN(.05) POUT(.05) ITERATE(20) .


*In the above, T_ assumes the same value as the survival time
indicator (SrvTmDg).
*'rik' assumes a value of 0 if the survival time is shorter
*than or equal to the time to decompensation (TmToRiks) and a
*value of 1 if the survival time is greater than TmToRiks (as
*this means that decompensation occurred, from the way the
*variables in the file were set up).
*'cancer' assumes a value of 0 if the survival time is shorter
*than or equal to the time to HCC (TmDgHCC) and a value of 1
*if the survival time is greater than TmDgHCC (as this means
*that HCC occurred, from the way the variables in the file
*were set up).

CLEAR TIME PROGRAM.
```

7.4 APPENDIX D. Calculating the empirical Brier score using a Cox model.

For any Cox model, the survival curve can be estimated for each combination of covariates based on the estimated baseline survival function and estimated model coefficients, so each individual in the sample may have a different estimated survival function $\hat{\pi}(t^*/X_j)$.

It is known that $S(t) = [S_0(t)]^{e^{\left(\sum_{j=1}^{p} \beta_j x_j\right)}}$ and essentially $\hat{\pi}$=S(t*).

For any particular Cox model, SPSS provides an estimated baseline survivor function $S_0(t)$ for each time point t in a separate file (NB ties are assigned slightly different values). Alternatively, the formula $S_0(t)=\exp(-H_0(t))$ can be used.

The following steps provide a detailed description of the calculations that can be performed to obtain the Brier score.

1. In excel, calculate $\beta^T x$ for each individual using the estimated β's from the Cox model. Obtain a column $\beta^T x$ next to columns of x's (creating dummy variables where necessary) with each individual 1,2,…,n in each ro,w as with the initial data set.
2. Call $\beta^T x$ the prognostic index, PI.
3. For each subject, exponentiate PI.
4. Insert T columns of constants $S_0(t)$, one for each chosen t*, t*=1,…,T e.g. t=12,24,36,48,60,72. Raise the estimated baseline survivor function $S_0(t^*)$ to the value obtained in step 3 for each subject. This gives the estimated S(t) for each subject, in row form.

To calculate the empirical Brier score, use the equation provided in **Section 2.1.4.1.** (Graf et al, 1999) for different t*:

1. Get the Kaplan-Meier (KM) censoring distribution G
2. Choose a t*
3. For each individual, see if their time in the study t≤t* & censored => not included in the calculation
4. If t≤t* & died => take $\hat{\pi}$ * $\hat{\pi}$* (1/ G at time of death)
5. If t>t* => take (1- $\hat{\pi}$)^2 * (1/ G at t*)
6. sum over individuals

7. divide by the number of individuals (including all censored individuals)

To get measure of explained residual variation:

1. Get the KM estimate at t* (for all patients) & call this $\bar{\pi}$ (t*)
2. Repeat the above steps 3-7
3. Use formula to calculate R^2.

This will give an estimate of the gain in accuracy in using the Cox model at each t*, as compared to the KM estimate.

7.5 APPENDIX E. S-PLUS: the bootstrap and jackknife-after-bootstrap techniques

I) The S-PLUS bootstrap function

```
> bootstrap
function(data, statistic, B = 1000, args.stat = NULL, group = NULL,
       sampler = samp.boot.mc, seed = .Random.seed, sampler.setup,
       sampler.wrapup, block.size = min(100, B), trace = T,
       assign.frame1 = F, save.indices = F, statistic.is.random,
       seed.statistic = 500)
{
# Capture call.
       func.call <- match.call()
       # Record unevaluated data and statistic as in the call.
       substitute.stat <- substitute(statistic)
       substitute.data <- substitute(data)
       # If statistic isn't function, store it as a call object to pass to
fit.func.
       if(mode(substitute.stat) == "call" || mode(substitute.stat) ==
              "{") statistic <- substitute.stat# Get name of data.
       data.name <- ifelse(length(substitute.data) == 1, deparse(
              substitute.data), "data")
       # Coerce vector to matrix so can index successfully.
       if(is.null(dim(data))) data <- as.matrix(data)
       # Get function to evaluate the statistic given data and indices.
       is.df.data <- is.data.frame(data)
       fit.func <- resamp.get.fit.func(statistic, substitute.stat,
              data.name, is.df.data, is.null(args.stat),
              assign.frame1)
       # Set seed in case statistic uses randomization
       seed <- eval(seed)
       if(missing(statistic.is.random)) {
              set.seed(seed.statistic)
              prev.seed <- .Random.seed
       }
#  Get parameter values for observed data.
       if(assign.frame1)
              on.exit(if(exists(data.name, frame = 1)) remove(
                            data.name, frame = 1))
       n <- dim(data)[1]
       observed <- fit.func(1:n, data, statistic, args.stat)
       # Determine if statistic uses randomization; this may fail if
#  a statistic sometimes use randomization.
       if(missing(statistic.is.random))
              statistic.is.random <- any(.Random.seed != prev.seed)
       if(statistic.is.random) seed.statistic <- .Random.seed
       # Check that observed is vector or matrix.  The need for a vector
# or matrix arises due to the use of apply to return a vector or vectorized
matrix.
       if(is.null(observed))
              stop("Statistic returned a NULL result on observed data.
  It must return a vector or matrix."
                     )
       if(!is.atomic(observed)) stop(
                     "Statistic must return a vector or matrix.")
       # Getting parameter names and coercing matrix to vector.
       names.observed <- resamp.get.dimnames(observed, substitute.stat
              )
       dim.obs <- dim(observed)
       if(!is.null(dim.obs))
              observed <- as.vector(observed)
       names(observed) <- names.observed# Sampler setup
       if(missing(sampler.setup))
              sampler.setup <- function(seed = 0)
              {
                     if(length(seed) == 1)
```

```
                        set.seed(seed)
                    else if(length(seed) == 12)
                        .Random.seed <<- seed
                    else stop("wrong seed length in sampler.setup")
                    return(seed)
                }
        if(missing(sampler.wrapup))
                sampler.wrapup <- function()
                return(.Random.seed)
        seed.start <- sampler.setup(seed)
        must.swap <- statistic.is.random & any(.Random.seed !=
                seed.statistic)
        # Need to swap only if both the sampler and statistic use .Random.seed
        call.stat <- function(i, fit.func, data, statistic, args.stat,
                inds.mat)
        fit.func(inds.mat[, i], data, statistic, args.stat)
        if(!missing(group)) {
# Find group using model.frame() stuff when have data frame.
# Note this doesn't apply for matrix or vector.
                if(is.df.data) {
                    m <- list(as.name("model.frame.default"), data
                            = func.call$data, group = func.call$
                            group)
                    mode(m) <- "call"
                    m <- eval(m, sys.parent())
                    group <- model.extract(m, group)
                }
# Get indices.
                group.inds <- split(1:n, group)
                ngroup <- length(group.inds)
        }
        nblocks <- ceiling(B/block.size)
        reps <- matrix(NA, length(observed), B)
        temp <- 1:block.size
        B2 <- block.size
        inds.mat <- matrix(NA, n, B2)
        if(save.indices)
                all.indices <- matrix(as.integer(0), n, B)
        on.exit({
                if(!all(is.na(reps))) {
                    B <- (i - 1) * block.size
                    cat("\nDid ", B,
                            " replications, saving results in .boots
trap.partial.results, interrupt again to abort completely.\n"
                            )
                    reps <- t(reps[, 1:B, drop = F])
                    dimnames(reps) <- list(NULL, names.observed)
                    func.call$B <- B
                    seed.end <- "Unknown, due to interrupt"
                    assign(".bootstrap.partial.results", where = 1,
                            immediate = T, bootstats(replicates =
                            reps, observed = observed, n = n, call
                             = func.call, seed.start = seed.start,
                            seed.end = seed.end, dim.obs = dim.obs,
                            group = group, indices = switch(
                              save.indices,
                            all.indices,
                            NULL)))
                }
        }
        , add = T)
        for(i in 1:nblocks) {
                if(trace)
                    cat("Forming replications ", 1 + (i - 1) *
                            block.size, " to ", min(i * block.size,
                            B), "\n")
                if(i == nblocks)
                    if(B %% block.size) {
```

```
                        B2 <- B %% block.size
                        temp <- temp[1:B2]
                        inds.mat <- inds.mat[, temp]
                }
        if(missing(group))
                inds.mat[] <- sampler(1:n, B2)
        else for(si in 1:ngroup)
                        inds.mat[group.inds[[si]],  ] <-
                          sampler(group.inds[[si]], B2)
        if(must.swap) {
                seed.sampler <- .Random.seed
                .Random.seed <<- seed.statistic
        }
        reps[, temp + block.size * (i - 1)] <- unlist(lapply(
                temp, call.stat, fit.func, data, statistic,
                args.stat, inds.mat))
        if(save.indices)
                all.indices[, temp + block.size * (i - 1)] <-
                        inds.mat
        if(must.swap) {
                seed.statistic <- .Random.seed
                .Random.seed <<- seed.sampler
        }
    }
    reps <- t(reps)     # Assign dimnames
    dimnames(reps) <- list(NULL, names.observed)
    seed.end <- sampler.wrapup()
    if(assign.frame1)
            on.exit(if(exists(data.name, frame = 1)) remove(
                        data.name, frame = 1))
    else on.exit()
    if(trace)
            cat("\n")
    bootstats(replicates = reps, observed = observed, n = n, call
             = func.call, seed.start = seed.start, seed.end =
            seed.end, dim.obs = dim.obs, group = group, indices =
            switch(save.indices,
            all.indices,
            NULL))
}
```

II) The S-PLUS jackknife after bootstrap function

```
> jack.after.bootstrap
function(boot.obj, functional = mean, threshold = 2, ..., frame.eval.boot =
sys.parent(1))
{
# Performs jackknife-after-bootstrap to obtain information on some
functional of
# the bootstrap distribution.  Returns estimates of the functional, its
standard
# error, and measures of the influence of each observation.  The standard
error
# estimates tend to be too large.  I'm interested in finding a well-
supported
# alternative, probably involving weighting.
# Hardwired options functional="Bias", "Mean", "SE".  Otherwise functional
is a function.
       if(!inherits(boot.obj, "bootstrap")) stop("boot.obj must be a
'bootstrap' object.")
       func.call <- match.call()
       func.call$functional <- substitute(functional)
       B <- boot.obj$B
       n <- boot.obj$n
       inds <- 1:n
       n.param <- length(boot.obj$obs)
# Get functional corresponding to "Mean", "Bias", or "SE".
```

```
	if(is.character(functional)) {
		if(functional == "Mean" || functional == "mean")
			functional <- mean
		else if(functional == "Bias" || functional == "bias")
			functional <- mean
		else if(functional == "SE" || functional == "se")
			functional <- function(x)
			{
				sqrt(var(x))
			}
		else stop("Functional must be a function or a character string
'Bias', 'Mean', or 'SE'.")
	}
# Get resampling indices.
	inds.mat <- resamp.get.indices(boot.obj, frame.eval.boot)
# Functional of full sample.
	func.full <- apply(boot.obj$rep, 2, functional, ...)
# Locate matches.
	has.match <- function(samp, target)
	duplicated(c(samp, target))[(length(samp) + 1):(length(samp) +
length(target))]
	matches.mat <- apply(inds.mat, 2, has.match, inds)
# Allocate space.
	func.vals <- matrix(nrow = n, ncol = n.param)
# Loop over parameters.
#* Maybe also calculate and store the mean of each subset of reps.
	jack.boot <- function(in.samp, reps, func, ...)
	{
		func(reps[!in.samp], ...)
	}
	for(j in 1:length(boot.obj$obs)) {
		func.vals[, j] <- apply(matches.mat, 1, jack.boot,
boot.obj$rep[, j], functional, ...)
	}
# Corrections if functional is "Bias".
	if(is.character(func.call$functional) && (func.call$functional ==
"Bias" || func.call$functional == "bias")) {
		func.full <- func.full - boot.obj$obs
		func.vals <- sweep(func.vals, 2, boot.obj$obs)
	}
# Calculate the SE(s) of the functional.
	if(any(is.na(func.vals)))
		stop("At least one observation is in every sample, so we cannot
calculate its influence.  Increase B and try again.")
	func.se <- apply(func.vals, 2, function(x, n)
	sqrt(((n - 1)/n) * sum((x - mean(x))^2)), n)
# Calculate jackknife influence values.
	rel.influence <- ( - (n - 1)) * scale(func.vals, center = T, scale =
sqrt(n) * func.se)  # Fiddle with names.
	names(func.se) <- names(boot.obj$obs)
	dimnames(func.vals) <- list(inds, names(boot.obj$obs))
	dimnames(rel.influence) <- dimnames(func.vals)
# Summary of relative influences.
	lri.func <- function(x, rel.inf, thresh)
	rel.inf[abs(rel.inf[, x]) >= thresh, x, drop = F]
	large.rel.influence <- lapply(names(boot.obj$obs), lri.func,
rel.influence, threshold)
	names(large.rel.influence) <- names(func.se)   # Return results.
	result <- list(call = func.call, functional = data.frame(Func =
func.full, SE.Func = func.se), rel.influence = rel.influence,
		large.rel.influence = large.rel.influence, values.functional =
func.vals, dim.obs = boot.obj$dim.obs, threshold = threshold)
	class(result) <- "jack.after.bootstrap"
	result
}
>
```

III) S-PLUS output for resampling procedures applied to the Cox model regression coefficients for the time to decompensation model

```
> attach(Tab2Clin1)

> boot.coxtab3<-bootstrap(Tab2Clin1,
+ coef(coxph(Surv(TmToRiks,Riksi)~AgeDiag+alc+hbv+hcv+alcvrs,
+ Tab2Clin1,na.action=na.omit)),B=1000,seed=0,trace=F)
)

> summary(boot.coxtab3)
Call:
bootstrap(data = Tab2Clin1, statistic = coef(coxph(Surv(TmToRiks, Riksi
       ) ~ AgeDiag + alc + hbv + hcv + alcvrs, Tab2Clin1, na.action =
       na.omit)), B = 1000, seed = 0, trace = F)

Number of Replications: 1000

Summary Statistics:
        Observed       Bias    Mean       SE
AgeDiag  0.01598  0.0007215  0.0167 0.007744
    alc  0.54365  0.0213268  0.5650 0.310593
    hbv  0.39984  0.0069000  0.4067 0.311125
    hcv -0.54175 -0.0005061 -0.5423 0.295861
 alcvrs  0.45444 -0.0301293  0.4243 0.470241

Empirical Percentiles:
             2.5%        5%      95%   97.5%
AgeDiag  0.002303  0.004425  0.02932 0.03237
    alc -0.022494  0.068496  1.08109 1.17384
    hbv -0.193405 -0.121942  0.91885 1.02227
    hcv -1.094236 -1.013532 -0.04805 0.05285
 alcvrs -0.587220 -0.341516  1.15461 1.27829

BCa Percentiles:
              2.5%        5%      95%   97.5%
AgeDiag  0.0003073  0.003288  0.02756 0.03093
    alc -0.0817291  0.003166  1.03515 1.12263
    hbv -0.2183393 -0.138510  0.89325 1.00015
    hcv -1.0768208 -1.000785 -0.01361 0.07881
 alcvrs -0.5449295 -0.309355  1.17658 1.32242

Correlation of Replicates:
        AgeDiag    alc    hbv    hcv alcvrs
AgeDiag  1.0000 0.1282 0.1039 0.1116 0.2309
    alc  0.1282 1.0000 0.6962 0.7278 0.4583
    hbv  0.1039 0.6962 1.0000 0.7129 0.4332
    hcv  0.1116 0.7278 0.7129 1.0000 0.4484
 alcvrs  0.2309 0.4583 0.4332 0.4484 1.0000

> limits.emp(boot.coxtab3, probs=c(0.025, 0.05, 0.5, 0.95, 0.975))
               2.5%           5%          50%         95%      97.5%
AgeDiag  0.002302595  0.004425374  0.01661882  0.02931801 0.03237126
    alc -0.022493575  0.068496041  0.56049002  1.08108630 1.17383837
    hbv -0.193405365 -0.121942083  0.40167593  0.91884596 1.02227156
    hcv -1.094236478 -1.013532089 -0.55318683 -0.04804968 0.05284853
 alcvrs -0.587220227 -0.341516028  0.44297274  1.15461232 1.27829000

 limits.bca(boot.coxtab3, probs=c(0.025, 0.05, 0.5, 0.95, 0.975),detail=T)
$limits:
                2.5%           5%          50%         95%      97.5%
AgeDiag  0.0003073356  0.003288221  0.01530439  0.02755938 0.03093014
    alc -0.0817291484  0.003165798  0.52709243  1.03515402 1.12262520
    hbv -0.2183393013 -0.138509977  0.39684600  0.89325327 1.00014673
    hcv -1.0768208306 -1.000784834 -0.53485325 -0.01360901 0.07881200
 alcvrs -0.5449294910 -0.309354536  0.47381791  1.17658228 1.32241916
```

```
> plot(boot.coxtab3)

> jab.coxtab3bias<-jack.after.bootstrap(boot.coxtab3,"bias")
> jab.coxtab3bias
Call:
jack.after.bootstrap(boot.obj = boot.coxtab3, functional = "bias")

Functional Under Consideration:
[1] "bias"

Functional of Bootstrap Distribution of Parameters:
             Func  SE.Func
AgeDiag  0.0007215 0.009808
    alc  0.0213268 0.396590
    hbv  0.0069000 0.403536
    hcv -0.0005061 0.382169
 alcvrs -0.0301293 0.572065

Observations with Large Influence on Functional:
```

$AgeDiag:

	AgeDiag
7	-2.494
15	-2.258
23	-4.978
37	-2.729
41	-2.998
45	-2.049
77	-2.427
79	2.318
82	2.786
91	-2.557
106	3.248
107	-2.083
108	2.413
115	2.729
130	2.279
183	-2.070

$alc:

	alc
1	-2.917
4	3.186
6	2.093
7	2.735
15	-3.015
17	-2.387
21	-2.083
25	4.385
32	-3.083
33	-2.748
37	-2.808
38	-3.261
41	-2.411
56	-2.034
75	-2.113
77	-3.221
93	-2.380
95	-2.127
295	-2.082

$hbv:

	hbv
1	-2.808
4	2.787
7	2.880
14	-2.494
15	-2.604
24	-2.010
25	3.758
31	-2.035
32	-2.759
33	-2.800
37	-2.034
38	-2.489
39	2.047
41	-2.719
108	2.083
109	-3.164
115	-2.198
134	2.523
140	-2.323
164	-2.117
275	2.082

$hcv:

	hcv
1	-2.953
4	3.615
7	2.955
14	-2.386
15	-2.461
21	-2.093
25	4.687
32	-2.708
33	-3.389
37	-2.342
38	-2.674
41	-2.348
164	-2.205
265	-2.086

$alcvrs:

	alcvrs
1	-2.439
17	-2.074
25	2.981
38	-2.537
253	-2.019
296	2.030
297	3.842
298	4.051
299	2.898
300	2.440
301	3.130
302	-2.294
304	-3.012
305	-5.369
306	-2.343
310	4.519
312	2.165

```
> plot(jab.coxtab3bias)
```

7.6 APPENDIX F. Calculation of a prognostic index in SPSS

```
*The mean age is subtracted from each individual age.

COMPUTE agediag0 = agediag - 62.29 .
VARIABLE LABELS agediag0 'mean 62.29 subtrctd' .
EXECUTE .

COXREG
  tmtoriks  /STATUS=riksi(1)
  /CONTRAST (gp2)=Simple(1)  /CONTRAST (gp3)=Simple(1)
/CONTRAST
  (gp4)=Simple(1)  /CONTRAST (gp5)=Simple(1)
  /METHOD=ENTER agediag0 gp2 gp3 gp4 gp5
  /SAVE=SURVIVAL HAZARD XBETA
  /PRINT=CI(95) BASELINE
  /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) .

*The PI calculated using the coefficients provided by the
*model.
*gp2 is a binary variable taking value 1 if individual had
*value 2 in the cirrhosis aetiology variable (ie alcohol)
*otherwise 0, gp 3 has value 1 if *HBV otherwise 0 and so on.

COMPUTE PI = (0.0158*agediag0) + (0.5417*gp2)+(0.3951 * gp3)-
(0.5372 *gp4)+(0.4523 * gp5) .
EXECUTE .
```

*The PI was exponentiated to obtain S(t).

```
COMPUTE expPI = EXP(pi) .
EXECUTE .

*Survivor functions were calculated for each subjects at 36
*months, 60 months & 84 months using the baseline survivor
*functions S0(36)=0.5226, S0(60)=0.3287, S0(84)=0.1836 .

COMPUTE surv3yr = 0.5226 ** expPI .
EXECUTE .
COMPUTE surv5yr = 0.3287 ** expPI .
EXECUTE .
COMPUTE surv7yr = 0.1836 ** expPI .
EXECUTE .
```

7.7 APPENDIX G. ENGLISH-GREEK GLOSSARY

A

absolute risk reduction, ARR	μείωση απολύτου κινδύνου, ΜΑΚ
accrual	προσαγωγή ή προσέλευση
actuarial estimator	αναλογιστικός εκτιμητής
albumin	λευκωματίνη
analysis of variance (ANOVA)	ανάλυση διασποράς
ANOVA table	πίνακας ανάλυσης διασποράς
antigen	αντιγόνο
azathioprine	αζαθειοπρίνη

B

backward elimination	επιλογή διαδοχικής αφαίρεσης
baseline hazard	βασική επικινδυνότητα
bed-side application	εφαρμογή κλίνης
bell-shaped	κωδωνοειδής
bias	μεροληψία
bilirubin	χολερυθρίνη
bimodal distribution	δικόρυφη κατανομή
binary data	δυαδικά δεδομένα
binary variable technique	διμεταβλητή τεχνική

C

categorical variable	ταξινομημένη μεταβλητή
censored	λογοκριμένο
censoring	λογοκρισία
chi-squared test	δοκιμασία χ^2 ή έλεγχος χ^2
clinical trial	κλινική δοκιμή
cohort	κοόρτη
compensated cirrhosis	αντιρροπούμενη κίρρωση
confidence interval (C.I.)	διάστημα εμπιστοσύνης (Δ.Ε.)
contingency table	πίνακας συνάφειας
continuous data	συνεχή δεδομένα
contrast	αντιπαραβολή, αντιπαράθεση,
control group	ομάδα ελέγχου
correlation	συσχέτιση
correlation coefficient	συντελεστής συσχέτισης
Cox proportional hazards (PH) regression model	μοντέλο αναλογικών επικινδυνοτήτων (ΑΕ) του Cox

critical values	κρίσιμες τιμές
cross-validation	διασταυρωτική επικύρωση
crude	αδρός
crude rate	αδρός ρυθμός
cumulative distribution function	συνάρτηση αθροιστικής κατανομής
cumulative hazard function	αθροιστική συνάρτηση επικινδυνότητας
cumulative incidence rate	αθροιστικός ρυθμός επίπτωσης
cumulative relative frequency	αθροιστική σχετική συχνότητα

D

data consistency	συμβιβαστότητα των δεδομένων
decompensation	ρήξη της αντιστάθμισης
decompensated cirrhosis	μη-αντιρροπούμενη κίρρωση
degrees of freedom, d.f.	βαθμοί ελευθερίας, β.ε.
density function	συνάρτηση πυκνότητας
deviation	απόκλιση
discrete	διακριτός
discrete data	ασυνεχής δεδομένα
discrete variable	διακριτή ή ασυνεχής μεταβλητή
discriminant analysis	διακρίνουσα ανάλυση
double blind trial	αμφιτυφλή δοκιμή
dummy variable	ψευδο-μεταβλητή

E

empirical Brier score	εμπειρική βαθμολόγηση Brier
end-point	τελικόν άκρον
error	σφάλμα
estimate	εκτίμιση
estimator	εκτιμητής ή εκτιμήτρια
event	ενδεχόμενο, συμβάν
event-free	ελεύθερος συμβάντων
evidence	τεκμήρια
explanatory variable	επεξηγηματική μεταβλητή
exponential distribution	εκθετική κατανομή

F

failure	αποτυχία
failure time	χρόνος αποτυχίας
follow-up time	χρόνος παρακολούθησης
forward selection	επιλογή διαδοχικής ένταξης

frame	πλαίσιο
frequency distribution	κατανομή συχνοτήτων
G	
goodness-of-fit	καλή ή όχι εφαρμογή
graph	γράφημα
grouped data	ομαδοποιημένα δεδομένα
H	
hazard	επικινδυνότητα
hazard function	συνάρτηση επικινδυνότητας
hazard ratio	λόγος επικινδυνότητας
hepatocellular	ηπατοκυτταρικό
hepatoma	ηπάτωμα
hepatitis B virus, HBV	ιος της ηπατίτιδας Β, HBV
hepatitis C virus, HCV	ιος της ηπατίτιδας C, HCV
histogram	ιστόγραμμα
I	
incidence	επίπτωση
independent censoring	ανεξάρτητη λογοκρισία
independent, identically distributed (iid) random variables	ανεξάρτητες, ταυτοτικά κατανεμόμενες (ατκ) τυχαίες μεταβλητές
indicator variable, Ind{}	δείκτρια, Δκτ{}
infection with HBV	μόλυνση με HBV
inference	συναγωγή συμπερασμάτων, συμπερασματολογία
informative censoring	πληροφοριακή λογοκρισία
instantaneous failure rate	στιγμιαίος ρυθμός αποτυχίας
interaction term	όρος αλληλεπίδρασης
intrinsic	εγγενής
L	
lack-of-memory property	ιδιότητα της έλλειψης-μνήμης
life expectancy	προσδόκιμο ζωής
life table	πίνακας επιβίωσης
likelihood ratio, LR	λόγος πιθανοφάνειας, ΛΠ
likelihoods	πιθανοφάνειες
limiting probability	ακραία πιθανότητα
line graph	γραμμογράφημα
log rank test	έλεγχος log rank ή δοκιμασία log rank

logistic regression	λογιστική παλινδρόμηση
loss-to-follow-up	χάσιμο (ασθενών) από παρακολούθηση
M	
masking	απόκρυψη
mean imputation	καταλογισμός του μέσου
median	διάμεσος
median follow-up time	διάμεσος του χρόνου παρακολούθησης
mortality rate	ρυθμός θνησιμότητας
N	
nested data	κιβωτισμένα δεδομένα
non-invasive techniques	μη επεμβατικές τεχνικές
non-parametric methods	μη παραμετρικές μέθοδοι
number needed to treat (NNT)	αριθμός ατόμων που απαιτείται να υποβληθούν σε θεραπεία (ΑΑΘ)
O	
octreotide	οκτρεοτίδη
ordinal data	διατάξιμα δεδομένα
outliers	ακραίες τιμές
overall	στο σύνολο, συνολικά
P	
paired	ανα ζεύγη
partition	διαμέριση
primary biliary cirrhosis, PBC	πρωτοπαθής χολική κίρρωση, ΠΧΚ
Pearson's correlation coefficient	συντελεστής συσχετίσεως του Pearson
person-years	ανθρωπο-έτη
placebo	πλασέμπο, εικονική θεραπεία
point estimation	σημειακή εκτίμηση
power law	δυναμοσυνάρτηση
power of a test	ισχύς μιας δοκιμασίας
predictive value	προγνωστική αξία
prevalance	επιπολασμός
probability density function, p.d.f.	πιθανοθεωρητική συνάρτηση πυκνότητας, π.σ.π
product limit estimator	εκτιμητής οριακού γινομένου
product limit method	μέθοδος οριακού γινομένου
prognostic index, PI	προγνωστικός δείκτης, Π.Δ.
progressively censored data	προδευτικώς λογοκριμένες παρατηρήσεις

prothrombin time ('quick' time)	χρόνος προθρομβίνης (χρόνος quick)
p-value	παρατηρούμενο επίπεδο σημαντικότητας ή τιμή p
R	
random variable	τυχαία μεταβλητή
randomized	τυχαιοποιημένο
randomized controlled trial	τυχαιοποιημένη ελεγχόμενη δοκιμή
range	εύρος ή πεδίο τιμών
rank sum test	αθροιστικός βαθμολογικός έλεγχος
ranked data	διατεταγμένα δεδομένα
rate	ρυθμός
ratio	λόγος
receptor	υποδοχέας
regression	παλινδρόμηση
regression coefficient	συντελεστής παλινδρόμησης
relative risk	σχετικός κίνδυνος
remaining lifetime	εναπομένουσα διάρκεια ζωής
replicates	επαναλαμβανόμενες
residual	υπόλοιπο
right censoring	δεξιόπλευρη λογοκρισία
risk test	κινδυνοσύνολο
S	
s.d.	τ.α.
sample size	μέγεθος δείγματος
sampling units	δειγματοληπτικές μονάδες
scatter diagram	στικτόγραμμα
score	βαθμός ή βαθμολογία
score test	δοκιμασία βαθμολογίας
score test statistic	κριτήριο της δοκιμασίας βαθμολογίας
SE or s.e.	ΤΑ ή τ.σ.
serum	ορός
serum protein electrophoresis	ηλεκτοφόρεσης πρωτεΐνης ορού
sign test	στατιστικός έλεγχος προσήμου ή δοκιμασία σημείων
significance level	επίπεδο σημαντικότητας
simple random sample	απλό τυχαίο δείγμα
simulation	προσομοίωση

singly censored data	μονόπλευρα λογοκριμένες παρατηρήσεις
skewed distribution	ασσύμετρη κατανομή
software package	λογισμικό πακέτο
somatostatin	σωματοστατίνη
Spearman's rank correlation coefficient	συντελεστής συσχέτισης διατάξεων του Spearman
standard deviation, s.d.	τυπική απόκλιση, τ.α.
standard error, S.E.	τυπικό σφάλμα, Τ.Σ.
standardised	προτυτοποιημένο
standardization	προτυτοποίηση
statistic (t, Z, X2)	κριτήριο (t, Z, X2)
statistical inference	επαγωγική στατιστική, στατιστική συμπερασματολογία
statistical significance	στατιστική σημαντικότητα
step function	βαθμιδωτή συνάρτηση
stepwise selection	επιλογή διαδοχικής ένταξης
study population	πληθυσμός υπό μελέτη
study units	μονάδες μελέτης
survival analysis	ανάλυση επιβίωσης
survival curve	καμπύλη επιβίωσης
survival function	συνάρτηση επιβίωσης
survival time	χρόνος επιβίωσης
T	
target population	πληθυσμός-στόχος
test	δοκιμασία ή έλεγχος
test statistic	κριτήριο της δοκιμάσιας
ties (in the data)	ισοβαθμίες (στα δεδομένα)
time decay effect	διαχρονική φθορά
time independent variable	χρονοανεξάρτητη μεταβληττή
time point	χρονική στιγμή
time to failure	χρόνος μέχρι την αποτυχία
time-dependent variable	χρονοεξαρτημένη μεταβληττή
time-fixed model	χρονοσταθερό μοντέλο
tree-based methods	μέθοδοι βασιζόμενες σε δενδροδιαγράμματα
U	
uncertainty	αβεβαιότητα
unimodal distribution	μονοκόρυφη κατανομή

univariate	μονομετάβλητη
V	
valid	έγκυρος
validate	επικυρώνω
validated	επικυρωμένος
variability	μεταβλητότητα
variance	διακύμανση
viral markers	ιολογικοί δείκτες
W	
withdrawal	απόσυρση

7.8 ΚΑΤΑΛΟΓΟΣ ΕΡΓΑΣΙΩΝ

Όνομα: Ιωάννα Μοσχανδρέα
Τίτλοι: BSc Mathematics (University of Sheffield, U.K.)
MSc in Applied Statistics (St. Hilda's College, University of Oxford, U.K.)
MSc in Quantitative Genetics & Genome Analysis (University of Edinburgh, U.K.)

Α. ΞΕΝΟΓΛΩΣΣΕΣ ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΧΕΤΙΚΕΣ ΜΕ ΤΗ ΔΙΑΤΡΙΒΗ

1. Samonakis DN, Moschandreas J, Arnaoutis T, Leontidis K, Skordilis P, Vafiades I, Kouroumalis E (in press, 2002) Treatment of hepatocellular carcinoma with long acting somatostatin analogues *Oncology Reports*

2. Koulentaki M, Ergazaki M, Moschandrea J, Spanoudakis S, Tzagarakis N, Drandakis PE, Spandidos DA, Kouroumalis EA (2001) Prevalence of hepatitis B and C markers in high-risk hospitalised patients in Crete: a five-year observational study. *BMC Public Health*;1(1):17

3. Alexandrakis MG, Moschandrea J, Kyriakou DS, Alexandraki R, Kouroumalis E (2001) Use of a variety of biological parameters in distinguishing cirrhotic from malignant ascites *International Journal of Biological Markers* 16 (1):45-49

4. Alexandrakis MG, Moschandrea J, Koulocheri SA, Kouroumalis E, Eliopoulos GD (2000) Discrimination between malignant and nonmalignant ascites using serum proteins in a multivariate analysis model *Dig Dis Sci* 45 (3): 500-8

5. Koulentaki M , Spanoudakis S, Kantidaki E, Drandakis P, Tzagarakis N, Biziagos E, *Moschandrea J*, Kouroumalis E.A. (1999) Prevalence of hepatitis B and C markers in volunteer blood donors in Crete – A five year study *J Viral Hepatitis* 6 (3): 243-8

6. Kouroumalis EA, Skordilis PG, Thermos K, Vasilaki A, Moschandrea J, Manousos O (1998) Treatment of hepatocellular carcinoma with octreotide. A randomized controlled study. *Gut* 42(3): 442-7

7. Kouroumalis EA, Skordilis PG, Moschandrea J, Alexandrakis G, Charoulakis N, Tzardi M, Manousos ON (1997) Natural history of advanced hepatocellular carcinoma in Crete and association with the hepatitis C virus. *Eur J Gastroenterology & Hepatology* 9 (10): 981-8

Β. ΑΛΛΕΣ ΞΕΝΟΓΛΩΣΣΕΣ ΔΗΜΟΣΙΕΥΣΕΙΣ

8. Moschandreas JA. & Kafatos A (in press, 2002) Calcium intake in relation to diet and health indicators in Cretan primary and high school pupils, Greece. *International Journal of Vitamin and Nutrition Research*.

9. Moschandreas J, Vissers MN, Wiseman S, van Putte K and Kafatos (in press, 2002) Extra virgin olive oil phenols and markers of oxidation in Greek smokers: a randomized crossover study *European Journal of Clinical Nutrition*

10. Manios Y, Moschandreas J, Hatzis C & Kafatos (in press, 2002) A health and nutrition education in primary schools of Crete: changes in chronic disease risk factors following a six year intervention programme. *British Journal of Nutrition*

11. Mouzas IA, Fragkiadakis N, Moschandrea J, Karachristos A, Skordilis P, Kouroumalis E & Manousos ON (2002) Validation and results of a questionnaire for functional bowel disorders in out-patients *BMC Public Health* 2(1):8

12. Skordilis P, Mouzas IA, Dimoulios PD, Alexandrakis G, Moschandrea J, Kouroumalis E (2002) Is endosonography an effective method for detection and local staging of the ampullary carcinoma? A prospective study. *BMC Surg* 25;2(1):1-7

13. Kafatos A, Verhagen H, Moschandreas J, Apostolaki I, van Westerop J.J.M. (2000) Mediterranean diet of Crete: foods and nutrient content *J Am Diet Assoc* 100(12):1487-93.

14. van de Vijver LP, Kardinaal AF, Couet C, Aro A, Kafatos A, Steingrimsdottir L, Amorim Cruz JA, Moreiras O, Becker W, van Amelsvoort JM, Vidal-Jessel S, Salminen I, Moschandreas J, Sigfusson N, Martins I, Carbajal A, Ytterfors A, Poppel G (2000) Association between trans fatty acid intake and cardiovascular risk factors in Europe: the TRANSFAIR study. *Eur J Clin Nutr* 54(2):126-35

15. Siest G, Bertrand P, Qin B, Herbeth B, Serot JM, Masana L, Ribalta J, Passmore AP, Evans A, Ferrari M, Franceschi M, Shepherd J, Cuchel M, Beisiegel U, Zuchowsky K, Rukavina AS, Sertic J, Stojanov M, Kostic V, Mitrevski A, Petrova V, Sass C, Merched A, Salonen JT, Tiret L, Visvikis S. and the ApoEurope group*. (2000) Apolipoprotein E polymorphism and serum concentration in Alzheimer's disease in nine European centres: the ApoEurope study. *Clin Chem Lab Med.* 38(8):721-30.
*Μέλος της ομάδας ApoEurope

16. Schiele F, De Bacquer D, Vincent-Viry M, Beisiegel U, Ehnholm C, Evans A, Kafatos A, Martins MC, Sans S, Sass C, Visvikis S, De Backer G, Siest G, The ApoEurope Group*. (2000) Apolipoprotein E serum concentration and polymorphism in six European countries: the ApoEurope Project. *Atherosclerosis.* 152(2):475-88.
*Μέλος της ομάδας ApoEurope

17. Moschandreas J & Kafatos A (1999) Food and nutrient intakes of Greek (Cretan) adults. Recent data for food based dietary guidelines *Br J Nutr* 81 *Suppl* 2: S71-6

18. Dontas AS, Moschandreas J, Kafatos A (1999) Physical activity and nutrition in older adults *Public Health Nutr* 2(3A):429-36

19. Hulshof KFAM, van Erp-Baart, Anttolainen M, Becker W, Church SM, Couet C, Hermann-Kunz E, Kesteloot H, Leth T, Martins I, Moreiras O, Moschandreas J, Pizzoferrato L, Rimestad AH, Thorgeirsdottir H, van Amelsvoort JMM, Aro A, Kafatos AG, Lanzmann-Petithory D, van Poppel G. (1999) Intake of fatty acids in Western Europe with emphasis on trans fatty acids: The TRANSFAIR study *Eur J Clin Nutr* 53,143-157

20. Manios Y, Moschandreas J, Hatzis C, Kafatos A (1999) Evaluation of a health and nutrition education program in primary school children of Crete over a three year period. *Prev Med* 28(2):149-59

21. Mouzas IA, Kremasmenou E, Kolivaki S, Giannadaki E, Skordilis P, Moschandrea J, Manousos ON (Abstract, 1996) High prevalence of perinuclear antineutrophil cytoplasmic antibodies in ulcerative colitis patients in Crete. *European Journal Of Internal Medicine*

Γ. ΕΛΛΗΝΙΚΕΣ ΔΗΜΟΣΙΕΥΣΕΙΣ

1. Μανιός Γ, *Μοσχανδρέα Ι,* Χατζής Χ, Καφάτος Α (2001) Παράγοντες κινδύνου χρόνιων νοσημάτων σε παιδιά δημοτικού της Κρήτης μετά την ολοκλήρωση εξαετούς προγραμματος αγωγής υγείας ΠΑΙΔΙΑΤΡΙΚΗ 65:66-76

Βραβείο "Κοινωνικής Παιδιατρικής" από την Ελληνική Παιδιατρική Εταιρεία στο 39^{ο} Πανελλήνιο Παιδιατρικό Συνέδριο / Κρήτη 2001

2. Λιναρδάκης, Μ, *Μοσχανδρέα Ι*, Καφάτος Α (2000) Καμπύλες σωματικής ανάπτυξης παιδιών βρεφικής και προσχολικής ηλικίας της Κρήτης που προέκυψαν από διαχρονική παρακολούθηση τους ΠΑΙΔΙΑΤΡΙΚΗ 63:391-407

Βραβείο "Κοινωνικής Παιδιατρικής" από την Ελληνική Παιδιατρική Εταιρεία στο 37^{ο} Πανελλήνιο Παιδιατρικό Συνέδριο / Θεσσαλονίκη 28-30.5.1999

3. Καφάτος Α, *Μοσχανδρέα Ι* , Χατζής Χ, Λιναρδάκης Μ, Βισβίκη Σ, Siest G, Ομάδα ApoEurope (2000) Οι πολυμορφισμοί της απολιποπρωτείνης Ε στην Κρήτη και η σχέση της με τη συγκέντρωση των λιποπρωτεϊνών και άλλους παράγοντες κινδύνου στεφανιαίας νόσου ΙΑΤΡΙΚΗ 77(5): 452-463

4. Καφάτος Α, Χατζής Χ, Μανιός Ι, Λιναρδάκης Μ, *Μοσχανδρέα Ι* , Μάρκατζη Ε (1999) Αγωγή υγείας στα σχολεία της Κρήτης. Ο τρόπος επεξεργασίας και παρουσίασης στους γονείς των κλινικών και εργαστηριακών αποτελεσμάτων των παιδιών τους ως μέρος της εκπαιδευτικής παρέμβασης ΠΑΙΔΙΑΤΡΙΚΗ 62:371-378

5. Καφάτος Α, Μανιός Γ, Χατζής Χ, *Μοσχανδρέα Ι*, Μπαλωμενάκη Ε, Αθανασόπουλος Δ, Μανιάς Ξ, Γαλανάκη Χ (1998) Αξιολόγηση προγράμματος «Αγωγή Υγείας μετά από τρία χρόνια εκπαιδευτικής παρέμβασης στα δημοτικά σχολεία της Κρήτης ΠΑΙΔΙΑΤΡΙΚΗ 61:483-497

6. Λιονής Χ, Μπαθιανάκη Μ, *Μοσχανδρέα Ι*, Βλαχονικολής Ι, Φιορέτος Μ Πρωτοβάθμια Φροντίδα και Ιατρείο Διαλογής σε Περιφεριακό Νοσοκομείο: Ο Ρόλος της Γενικής/Οικογενειακής Ιατρικής (1996) Πρωτοβάθμια Φροντίδα Υγείας, 8 (2): 77-86

Το άρθρο έλαβε το πρώτο βραβείο στο διαγωνισμό της Ελληνικής Εταιρείας Γενικής Ιατρικής εις μνήμη του καθηγητή Δ. Ίκκου, 8^{ο} Πανελλήνιο Συνέδριο Γενικής Ιατρικής, Χαλκιδική, 17-21 Απριλίου 1996

Δ. ΣΥΜΜΕΤΟΧΗ ΣΕ ΤΕΧΝΙΚΕΣ ΑΝΑΦΟΡΕΣ (TECHNICAL REPORTS)

1. TNO report V 98.593 Food consumption factors in relation to packaging materials in 8 European countries (1998) MRH Löwik, KFAM Hulshof, C Cuadrado, I Elmadfa, S. de Henauw, L Johansson, A Kafatos, O Moreiras, *J Moschandreas*, L Pizzoferrato, JL Volatier

2. Institute of European Food Studies (IEFS), Trinity College, Dublin (1999) A pan-EU survey on consumer attitudes to physical activity, body-weight and health Office for Official Publications of the European Communities, Luxembourg. Μέλος του Project Management Group.