Transpl Int (2003) 16: 794–800
DOI 10.1007/s00147-003-0625-z

ORIGINAL ARTICLE

Michael Heise
Utz Settmacher
Robert Pfitzmann
Ute Wünscher
Andrea Raffaela Müller
Sven Jonas
Peter Neuhaus

A survival-based scoring-system for initial graft function following orthotopic liver transplantation

Received: 8 October 2002
Revised: 18 February 2003
Accepted: 11 March 2003
Published online: 3 July 2003

M. Heise (✉) · U. Settmacher
R. Pfitzmann · U. Wünscher
A.R. Müller · S. Jonas · P. Neuhaus
Department of Surgery,
Humboldt University of Berlin, Charité,
Augustenburger Platz 1,13353 Berlin,
Germany
E-mail: michael.heise@charite.de
Tel.: +49-30-450552001
Fax: +49-30-450552000

Abstract Initial graft function following orthotopic liver transplantation is a major determinant of postoperative survival and morbidity. Despite several efforts to provide scoring-systems for initial graft function, there is still a lack of a generally accepted classification scheme. The previously published systems assessed initial graft function based on the first postoperative days or weeks using liver-related laboratory parameters. It was shown that in most cases the scoring-systems did not correlate with patient survival. We intended to refine the definition of initial graft function in order to provide a survival based classification system. In a retrospective analysis of 761 patients following primary liver transplantation, a new scoring-system for early postoperative graft function was developed. Statistically significant differences in long term survival were calculated for ALAT, ASAT, bile production and prothrombin activity on days 1, 3, 7, 14. Points were then assigned according to the degree of survival: improved survival = 1 point, poor survival = 2 points. Patients were split into three groups corresponding to initially good, moderate and poor function. Applying this score, early and late patient survival rates and incidence of initial non-function were statistically significantly different. This was in contrast to the Gonzalez and the Ploeg-Maring classification scales, which are based on arbitrarily chosen cutoff levels. Retransplantation rates and postoperative morbidity were comparable both for the new and the older systems. We can conclude that the presented refined scoring-system for initial graft function provides a significant correlation to patient survival and initial non-function. We recommend the refined system for future studies.

Keywords Initial non-function · Scoring systems · Liver transplantation · Patient outcome · Posttransplant Survival

Introduction

Despite efforts to expand the available donor pool, donor shortage remains a major problem in clinical transplantation [1]. Various studies are used to define donor risk factors for initial poor or non-function [2, 3, 4, 5, 6, 7]. In this context, various factors have been identified to correlate with postoperative graft function [3, 8, 9, 10]. However definitive assessment is difficult since most studies use differing endpoints and scoring-systems [3, 11]. Hence, the overall results are inconsistent. One of the main reasons is because the cutoff levels used to classify initial graft function are arbitrarily chosen, throughout the published studies, and not based on a statistically sound rationale. The previously published systems intended to asses initial graft function based on the different time intervals during the first 7 postoperative days, using liver-related laboratory parameters, such as transaminases levels and prothrombin activity. These parameters, however, are extremely variable from day to day during the first weeks, depending on graft recovery and presence of acute rejection. We therefore developed a simple classification system, based on transaminases levels, prothrombin activity and bile production on 4 postoperative days. In contrast to the known scoring-systems, cutoff levels were chosen with regard to improved or poor patient long term survival. Patients were split into three groups, and survival and postoperative morbidity rates were compared with those of the Gonzalez and Ploeg-Maring classification systems [3, 4, 5].

Material and methods

A retrospective analysis was performed on 761 patients who underwent orthotopic liver transplantation between 9/88 and 6/97. Only primary transplantations with use of University-of-Wisconsin-Solution for graft procurement were enrolled. Pediatric cases < 14 years were excluded. 316 (41.5%) female and 445 (58.5%) male patients at a mean age of 46 ± 11 years were enrolled. Indications for orthotopic liver transplantation are shown in Table 1.

Early (within 30 days) and long term patient survival served as end points. Postoperative morbidity was assessed according to the need for dialysis and reintubation, according to the incidence of pleural effusion and the rate of surgical revisions. Mean survival rates were compared using the Kaplan-Meier method and log rank test. For univariate analyses, variables were dichotomized and cutoff values identified. An univariate survival analysis using ALAT, ASAT, prothrombin activity and bile output per day was performed for postoperative days (POD) 1, 3, 7, 14 in order to identify statistically significant cutoff values within the individual variables. Points were assigned for the degree of survival (Table 2): improved survival = 1 point, poor survival = 2 points. Cases with cumulative scores of 4 were defined as class A (good function), 5–6 as class B (intermediate function) and 7–8 as class C (initial poor function). Patients were also grouped according to the preoperative Child score [12].

In addition, early postoperative graft function was assessed by means of the Ploeg score, refined by the Maring and the Gonzalez classification systems (Table 3) [3, 5]. All three definition systems

Table 1 Indications for orthotopic liver transplantations

Indications	Number of patients (n = 761)	Percentage
HBV-cirrhosis	134	17.6
HCV-cirrhosis	148	19.4
Alcoholic cirrhosis	139	18.3
PBC	72	9.5
PSC	35	4.6
HCC	28	3.7
Autoimmune hepatitis	25	3.3
Budd-Chiari-syndrome	16	2.1
Other	164	21.3

Table 2 The Berlin criteria used to define initial graft function. Cutoff levels for ASAT, ALAT, prothrombin activity and bile production were based on statistical significant differences in long-term survival (log-rank test). When significant differences were not determined, values corresponding to most diverging curves were used

Parameter	POD1			POD3			POD7			POD14		
	1 point	2 points	*P*-value	1 point	2 points	*P*-value	1 point	2 points	*P*-value	1 point	2 points	*P*-value
ASAT (U/l)	≤ 1800	> 1800	< 0.05	≤ 550	> 550	< 0.05	≤ 30	> 30	n.s.	≤ 50	> 50	< 0.05
ALAT (U/l)	≤ 1600	> 160	< 0.05	≤ 650	> 650	< 0.05	≤ 150	> 150	n.s.	≤ 150	> 150	< 0.05
Prothrombin activity %	≥35	< 35	< 0.05	≥45	< 45	< 0.05	≥45	< 45	< 0.05	≥70	< 70	< 0.05
Bile production (ml/d)	≥50	< 50	n.s.	≥50	< 50	< 0.01	≥200 or T-tube closed	< 200	< 0.05	≥200 or T-tube closed	< 200	< 0.05
Score	Σ4, 5 = initial good function (Berlin A), Σ6 = intermediate function (Berlin B), Σ7, 8 = initial poor function (Berlin C)											

Table 3 Gonzalez and Ploeg-Maring criteria used for calculation of initial graft function following liver transplantation. Within the Gonzalez system, grafts with a cumulative score of 3 or 4 points were assigned to those with good function, those with a score of 5 or 6 to those with moderate function, and those with 7 to 9 to the group with initial poor function *IPF*. *FFP* fresh-frozen plasma

Gonzalez				Ploeg-Maring		
Parameter	1 point	2 points	3 points	Parameter	No IPF	IPF
ALAT (U/l)	< 1000	1000–2500	> 2500	ASAT(U/l)	< 2000	> 2000
Prothrombin activity %	≥60	> 60 with FFP	< 60 despite FFP	Prothrombin time (s)	< 16 s	> 16 s
Bile production (ml/d)	> 100	40–100	< 40	Sample period	Between days 1–7	
Sample period	72 h after transplantation					

were eventually compared in terms of mean patient survival, early and late retransplantation rates, mean time to retransplantation and risk of initial non-function. Morbidity in terms of incidence of dialysis, reintubation, pleural effusion and surgical revisions was calculated and related to the classification systems. Initial non-function (INF) was defined as inability to support life, and necessitating early retransplantation. The results were evaluated by Kruskal-Wallis and $\chi 2$ tests. $P < 0.05$ were considered statistically significant.

Results

Initial graft function, survival and postoperative morbidity

Child score

109 patients (15.6%) were classified as Child A, 349 (50%) as Child B and 239 (34.4%) as Child C. The survival rates of patients grouped according to the preoperative Child score was statistically not significantly different (Fig. 1).

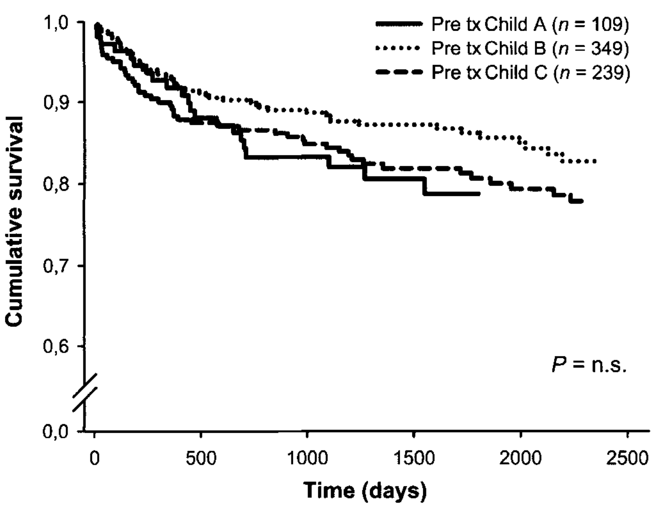


Fig. 1 Cumulative survival of recipients grouped according to their preoperative Child-Pugh scores

Berlin score

The patients were distributed evenly over the three groups during the first 2 postoperative weeks (Table 4). Between days 1 and 14, a mean of 674 ± 8 patients (89%) were classified as having good initial function (IGF, class A), 70 ± 5 (9%) as having intermediate function (IMF, class B), and 9 ± 5 (2%) as suffering from initial poor function (IPF, class C). Patient's migration between the individual classes during the postoperative period is depicted in Fig. 2. Early and late retransplantation rates, mean time to retransplantation, and mean patient survival were statistically significantly different between the groups (Table 3, Figs. 3, 4). Incidence of initial non-function (INF) was one case in group A (0.1%), 3 patients in group B (2%) and 3 (38%) in group C. Postoperative morbidity in terms of need for dialysis, reintubation and surgical revisions was statistically significantly different between the three groups (Table 5). Incidences of pleural effusion and tracheotomy were higher in the poor classes, but the differences were statistically not significant.

Gonzalez score

Applying the Gonzalez criteria on the first postoperative day, only 249 patients (33%) were classified as having good initial function (IGF, class I), while 387 (51%) were considered to have moderate function (IMF, class II), and 125 (16%) to be suffering from initial poor function (IPF, class II). Between postoperative days 7 and 14, 677 + 21 (90%) were classified as IGF, 68 + 23 (9%) patients with IMF and 4 + 2 (1%) as IPF. However, mean patient survival after 2600 days was statistically not significantly different (Fig. 3). Early and late retransplantation rates and the mean time to retransplantation differed significantly between the three groups (Table 4). Of all patients who were classified at POD 1, no patient of the IGF group, but 1 (0.3%) of the IMF and 10 (8%) of the IPF group developed initial non-function. Conversely, of all patients who were classified according to the Gonzalez criteria at POD 7, none of group I, but 2 (2.4%) and none of group III developed initial

Table 4 Comparison of graft and patient survival, retransplantation and initial non-function rates of patients grouped according to the Berlin, Gonzalez and Ploeg-Maring criteria. # = limited information due to low case numbers

	POD	Berlin Score				Gonzalez Score				Ploeg-Maring Score		
		A	B	C	*P*-value	I	II	III	*P*-value	I	II	*P*-value
Number of patients	1	686	63	12	< 0.01	249	387	125	< 0.01	706	55	< 0.01
	3	668	72	16	< 0.01	470	233	53	< 0.01	747	9	< 0.01
	7	673	75	3	< 0.01	662	84	5	< 0.01	749	2	< 0.01
Recipient survival per cent	1	82	71	80	< 0.05	82	79	77	n.s.	81	73	n.s.
	3	82	66	62	< 0.01	81	74	68	n.s.	78	44	n.s.
	7	84	69	60	< 0.01	81	71	80	n.s.	80	0	n.s.
Early retransplantation rate (< 30 days, per cent)	1	0,7	10	42	< 0.01	0,4	0,8	10	< 0.01	0,6	22	< 0.01
	3	0,6	4	25	< 0.01	0,6	4	25	< 0.01	0,5	78	< 0.01
	7	0,2	2	60	< 0.01	0,2	4	25	< 0.01	0,5	100	< 0.01#
Total retransplantation rate (< 2600 days per cent)	1	7	17	42	< 0.01	6	6	20	< 0.01	7	25	< 0.01
	3	7	14	31	< 0.01	7	14	31	< 0.05	7	89	< 0.01
	7	6	11	60	< 0.01	7	10	40	< 0.05	7	100	< 0.01
Mean time to retransplantation (days)	1	3586	2769	1927	< 0.01	3571	3623	3035	< 0.01	3582	2357	< 0.01
	3	3556	3014	2373	< 0.01	3559	3464	2835	< 0.01	3542	1615	< 0.01
	7	3639	3387	1444	< 0.01	3600	3244	1555	< 0.01	3586	9	< 0.01#
INF rate, %	1	0,1	8	42	< 0.01	0	0,3	8	< 0.01	0	20	< 0.01
	3	0	4	19	< 0.01	0	0	11	< 0.01	0	67	< 0.01#
	7	0	0,6	20	< 0.01	0	2	0	< 0.01	0	100	< 0.01#

non-function with a need for undergoing early retransplantation. Postoperative morbidity in terms of need for dialysis was significantly increased both for the IMF and IPF groups. The remaining morbidity factors, such as incidence of reintubation, tracheotomy, pleural effusion and surgical revisions increased proportionally with deteriorating groups, but were not in all instances statistically significant (Table 5).

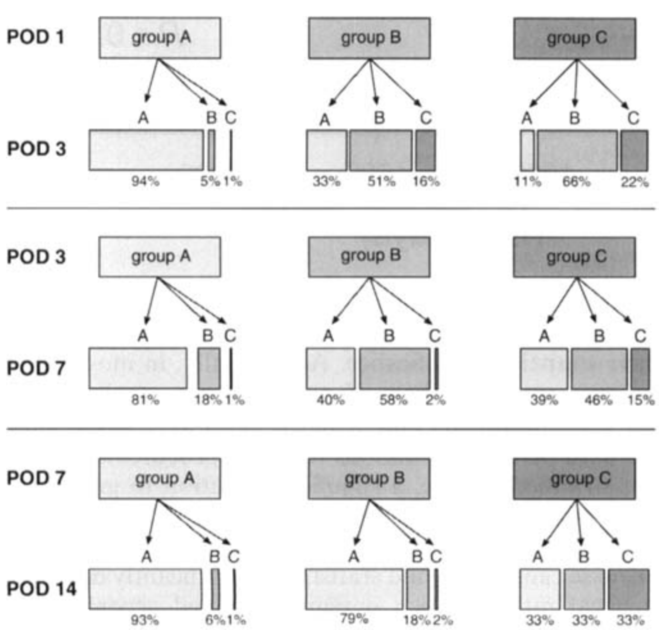


Fig. 2 Migration of patients classified by means of the Berlin criteria on POD one, three, seven and fourteen between the individual groups. While most patients remain in group A, significant proportions of group B and C patients change during the postoperative course

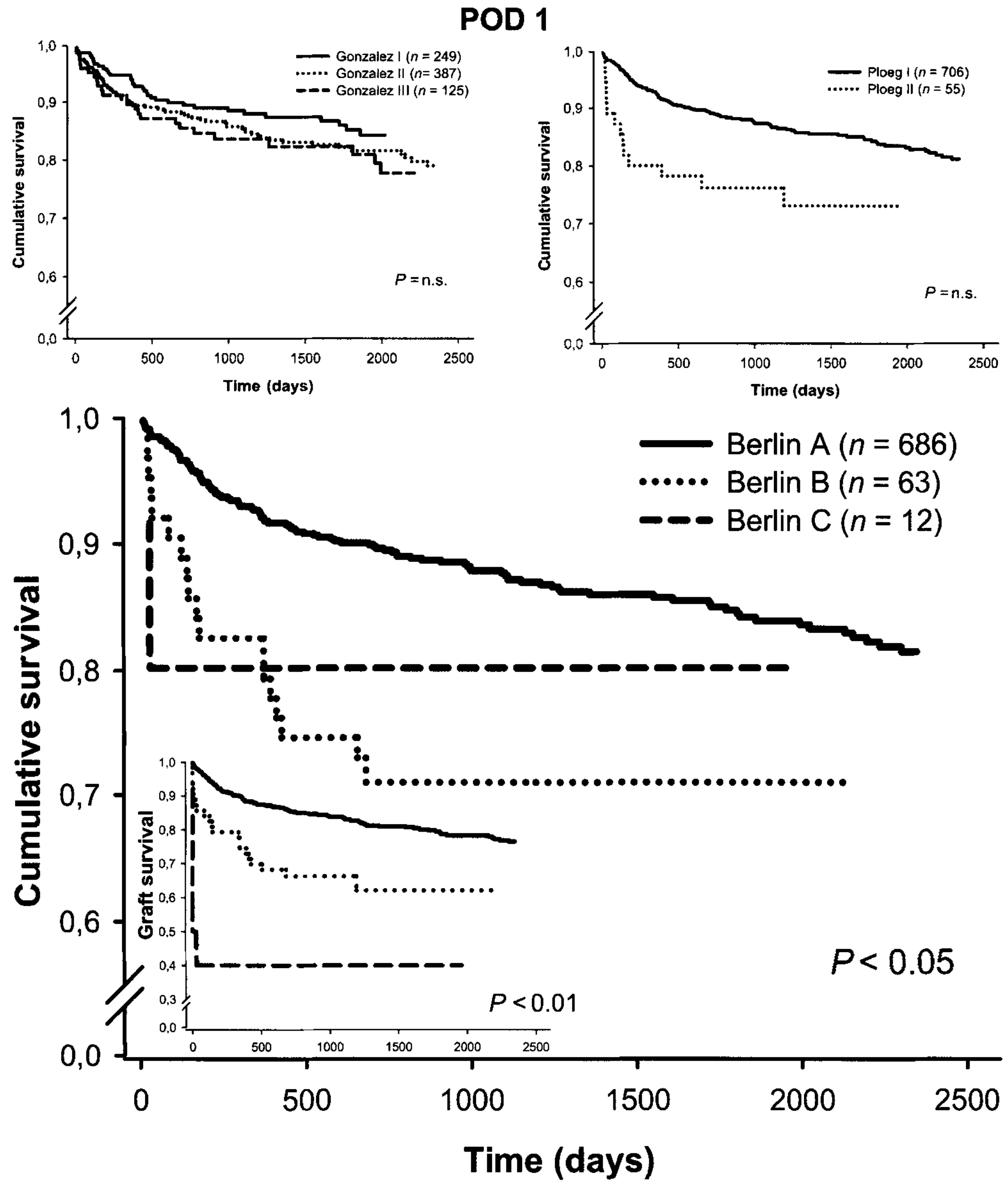
Ploeg-Maring score

According to the Ploeg-Maring classification, the vast majority of patients 706 (93%) were grouped on POD 1 as IGF (group I) and 7% as IPF (group II). On POD 7, 99.7% were grouped as IGF, and only 0.3% as IPF. Mean patient survival after 2600 days was statistically not significantly different on POD 1 (Fig. 3). On POD 7, the size of group II was too small ($n = 2$) for a valid estimation of survival by means of the Kaplan-Meier method. Early and late retransplantation rates as well as mean time to retransplantation differed significantly between both groups (Table 4). The incidence of initial non-function was 20% for the IPF group, while no case of INF was observed for the IGF group. Postoperative incidences of both dialysis and reintubation rates were significantly increased between group I and II. Although the rates of tracheotomy, pleural effusions and surgical revisions were higher in the IPF group, the differences were statistically not significant (Table 5).

Discussion

The first postoperative weeks following orthotopic liver transplantation are a dynamic period in terms of surgical convalescence, graft recovery and incidence of acute rejections [7, 13, 14]. Postoperative course and survival of the patients depends in the first place on the initial function of the transplanted liver [3, 15]. This is reflected

Fig. 3 Long-term survival curves of patients classified on POD one according to the Gonzalez, Ploeg-Maring and Berlin criteria for initial graft function. The long term survival rate of Berlin group C is superior than that of A and B due to a high rate of initial nonfunction with required subsequent retransplantation. This is underlined by the graft survival rates which are shown in the inset



by the fact that the postoperative survival did not correlate with the preoperative Child score.

During the last years, several attempts have been made to assess initial graft function, mostly in combination with evaluation of donor-specific influences [3, 4, 5, 10, 16, 17]. In most cases, liver specific parameters like transaminases levels, prothrombin time or activity and bile production have been used for the classification of hepatic grafts during the first postoperative days and weeks. Other studies included bilirubine, ammonia levels or postoperative encephalopathy [4, 18]. However, despite several efforts to refine the definition and classification scale of initial graft function, a widely respected system is not yet available. This may be due to the fact, that the applied variables in the mentioned studies were very heterogeneous in terms of applied cut-off levels and their statistical significance. Additionally, in most studies only small numbers of patients were enrolled, primary and secondary and pediatric cases were mixed, and the used perfusion solutions for graft procurement were not stratified. Hence, a significant relation to graft survival could not always be shown [3, 4, 5, 11]. This is confirmed by the results of the presented study, which likewise cannot not find statisticaly significantly different survival rates between patients classified according to the Gonzalez or Ploeg-Maring criteria. Since graft and patient survival remains the most important clinical aspect, we attempted to define a classification system which allows allocating the individual patient to a group that provides a significant and validated correlation with

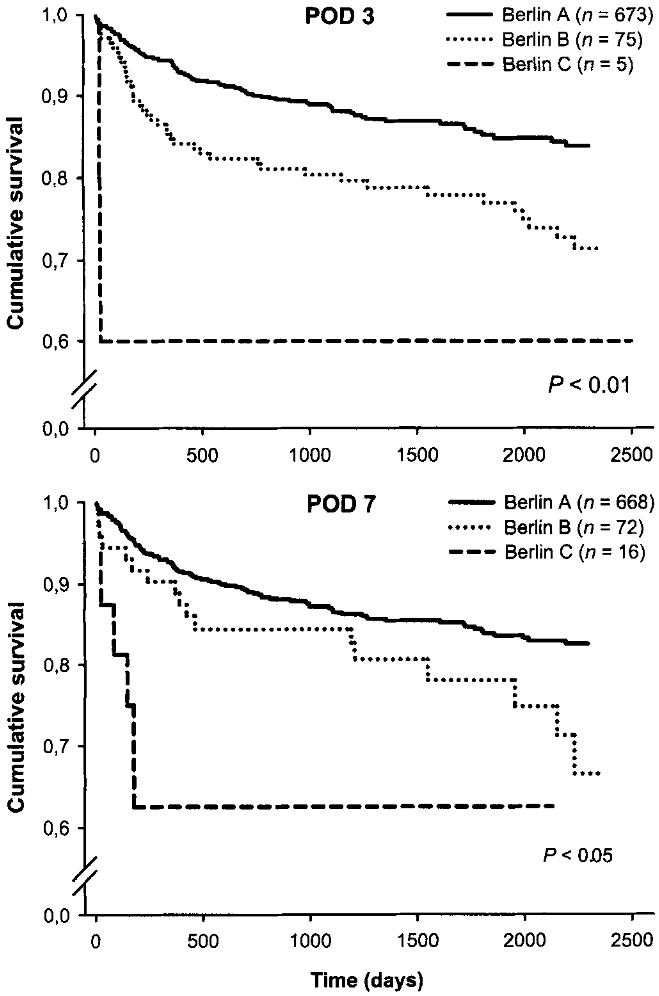


Fig. 4 Long-term survival curves of patients classified on POD 3, and 7 according to the Berlin score

long term survival. Two aspects were important. First, the value of the cutoff level had to be refined in terms of long-term survival, and, secondly, the dynamics of the postoperative course, with significantly declining laboratory values, had to be reflected. The known scoring-systems summarized the first postoperative days or the first week. Application of this imprecise classification patterns regardless of the dynamics of the postoperative laboratory values led to grouping patients rather randomly.

From initially high levels, a significant decline is seen in most of the patients. It therefore seems impossible to define initial graft function by laboratory values which were averaged over several days or weeks. This was confirmed by the fact that only few patients could be assigned to the IPF group of the Ploeg-Maring Score, which considers the entire first week [4]. In order to refine the known scoring-systems, it was essential to assess the graft function within shorter time intervals. We used four time points (POD 1, 3, 7, 14) for this purpose, which were chosen empirically according to clinical aspects.

The second important aspect addresses the numerical value of the cutoff levels applied to group the patients. The selected levels of the known systems were arbitrary and not based on a proven rationale. Although it appeared reasonable to define ASAT/ALAT levels higher than 1.000 or 1.500 U/l and prothrombin activity of less than 60% as cutoff levels for poor function, it would be even more reasonable if these levels would correlate with survival. We agree with Strasberg who questioned the ascertaining the gravity of IPF without any correlation to patient survival [11]. As shown in this study, the survival rates both in the Gonzalez and the Ploeg-Maring groups were not significantly different, with exception of the Gonzalez score for POD 3.

We intended to make a a simple scoring-system available that can provide a relation to long-term survival. Due to the mentioned variability of the lab values within the first postoperative days and weeks, it is not possible to define general cutoff levels, which are valid between days 1 and 14. Therefore, individual cutoff levels for selected days were determined, and in most cases significant cutoff levels could be identified. Only bile production on POD1 and the transaminases levels on POD 7 were statistically not significant. In these cases, the values corresponding to the most diverging

Table 5 Postoperative morbidity of patients grouped on day 1, 3, 7, 14 according to the Berlin, Gonzalez and Ploeg-Maring criteria

	POD	Berlin score				Gonzalez score				Ploeg-Maring score		
		A	B	C	*P*-value	I	II	III	*P*-value	I	II	*P*-value
Need for dialysis (per cent)	1	11	41	33	<0.01	8	14	28	<0.01	13	38	<0.01
	3	12	25	56	<0.01	11	16	34	<0.01	14	67	<0.01
	7	12	18	80	<0.01	12	29	60	<0.01	14	0	<0.01
Reintubation required (per cent)	1	6	11	25	<0.05	6	7	9	n.s.	6	16	<0.01
	3	6	7	25	<0.05	6	5	19	<0.01	6	56	<0.01
	7	6	7	40	<0.01	5	18	40	<0.01	6	100	<0.01
Surgical revisions required (per cent)	1	16	16	17	n.s.	11	17	21	<0.05	15	20	n.s.
	3	15	17	31	n.s.	13	20	23	<0.05	16	22	n.s.
	7	15	17	60	<0.05	14	20	100	<0.05	16	50	n.s.

survival curves were applied. Points were then assigned according to the degree of survival; improved survival = 1, and poor survival = 2 points. The cumulative scores allowed grouping the patients into three groups (Berlin A to C) with varying initial graft function and statistically significantly differing survival rates, with exception of the first day, on which the survival in group B was somewhat lower than that of group C. This was ascribed to the small number of patients in group C and the high percentage of INF cases. 42% of the Berlin C patients of POD 1 required an early retransplantation, resulting in subsequently improved survival for this particular group. However, the remaining survival proportion rates between day 3 and 14 correlated very well with worsening classes, providing a mean survival rate of 83% for group A and 60% for group C (2600 days). Initial non-function rates within the Gonzalez and Ploeg-Maring systems were significantly higher for the poor groups, but the latter provided better separation.

Statistically significant differences were also found for the Berlin and the Gonzalez systems in terms of postoperative morbidity. Patients grouped by means of the Gonzalez and the Berlin Score had a comparable postoperative morbidity. The incidences for dialysis, reintubation, pleural effusion and surgical revisions were significant different for most investigated time points. In this respect, both the presented and the Gonzalez scoring-systems were comparable. The Ploeg-Maring criteria could only be applied between day 1 and 3, since no patient presented with ASAT levels of <2000 U/l on day 7. They were either dead or had undergone retransplantation.

In times of decreasing donor rates the donor pool has to be expanded, and therefore marginal donors have to be accepted more frequently. The question whether the use of livers from marginal donors has a significant impact on postoperative mortality and morbidity has to be addressed in future studies. Since the known studies used variables with arbitrarily chosen cutoff levels to determine study endpoints, an improvement of the donor-recipient relationship seems possible.

References

1. Loinaz C, Gonzalez EM. Marginal donors in liver transplantation. Hepatogastroenterology 2000; 47: 256–263
2. Melendez HV, Rela M, Murphy G, Heaton N. Assessment of graft function before liver transplantation: quest for the lost ark? Transplantation 2000; 70: 560–565
3. Maring JK, Klompmaker IJ, Zwaveling JH, Kranenburg K, Ten Vergert EM, Slooff MJ. Poor initial graft function after orthotopic liver transplantation: can it be predicted and does it affect outcome? An analysis of 125 adult primary transplantations. Clin Transplant 1997; 11: 373–379
4. Ploeg RJ, D'Alessandro AM, Knechtle SJ et al. Risk factors for primary dysfunction after liver transplantation–a multivariate analysis. Transplantation 1993; 55: 807–813
5. Gonzalez FX, Rimola A, Grande L et al. Predictive factors of early postoperative graft function in human liver transplantation. Hepatology 1994; 20: 565–573
6. Gruenberger T, Sautner T, Wamser P, Mittlbock M, Steininger R, Muhlbacher F. Impact of donor criteria on postoperative graft function in liver transplantation. Transplant Proc 1993; 25: 3156–3157
7. Mueller AR, Platz KP, Krause P, Neuhaus R, Lobeck H, Neuhaus P. Does initial graft function influence the outcome after liver transplantation? Transplant Proc 1998; 30: 1464–1465
8. D'Alessandro AM, Kalayoglu M, Sollinger HW et al. The predictive value of donor liver biopsies on the development of primary nonfunction after orthotopic liver transplantation. Transplant Proc 1991; 23: 1536–1537
9. Kizilisik TA, al Sebayel M, Ramirez CB et al. Analysis of donor criteria and its implications on the outcome of clinical liver transplants. Transplant Proc 1997; 29: 2836–2838
10. Greig PD, Forster J, Superina RA et al. Donor-specific factors predict graft function following liver transplantation. Transplant Proc 1990; 22: 2072–2073
11. Strasberg SM, Howard TK, Molmenti EP, Hertl M. Selecting the donor liver: risk factors for poor function after orthotopic liver transplantation. Hepatology 1994; 20: 829–838
12. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg 1973; 60: 646–649
13. Mueller AR, Platz KP, Krause P et al. Perioperative factors influencing patient outcome after liver transplantation. Transpl Int [Suppl 1] 2000; 13: 158–161
14. Wood RP, Rosenlof LK, Shaw B-WJ, Pillen TJ, Williams L. Complications requiring operative intervention after orthotopic liver transplantation. Am J Surg 1988; 156: 513–518
15. Pokorny H, Gruenberger T, Soliman T, Rockenschaub S, Langle F, Steininger R. Organ survival after primary dysfunction of liver grafts in clinical orthotopic liver transplantation. Transpl Int 2000; 13 Suppl 1
16. Bein T, Frohlich D, Pomsl J, Forst H, Pratschke E. The predictive value of four scoring-systems in liver transplant recipients. Intensive Care Med 1995; 21: 32–37
17. Briceno J, Solorzano G, Pera C. A proposal for scoring marginal liver grafts. Transpl Int [Suppl 1] 2000; 13: 249–252
18. Deschenes M, Belle SH, Krom RA, Zetterman RK, Lake JR. Early allograft dysfunction after liver transplantation: a definition and predictors of outcome. National Institute of Diabetes and Digestive and Kidney Diseases Liver Transplantation Database. Transplantation 1998; 66: 302–310