

# Kapitel 3

## Ergebnisse

Insgesamt wurden 206 Patienten in die Studie aufgenommen, aber nur bei 193 Patienten konnte der Goldstandard der Dignität (benigne oder maligne Läsion) ermittelt werden. Sieben Patienten mussten ausgeschlossen werden, da von ihnen keine Follow-up-Daten vorlagen und weitere sechs, weil ihre Läsion außerhalb des Pankreas lag. 99 (51%) der 193 Patienten der Auswertung waren Männer, 94 (49%) Frauen. Das Durchschnittsalter betrug 61 Jahre, der jüngste Patient war 20, der älteste 87 Jahre alt.

Von den 193 Läsionen waren 119 (62%) maligne, hauptsächlich duktale Adenokarzinome, und 74 (38%) benigne (vgl. Tabelle 3.1). Die meisten Läsionen befanden sich im Pankreaskopf (85%) und waren zwischen 2 und 4 cm groß (siehe Tabelle 3.2).

Bei den 119 Patienten mit maligner Läsion erfolgte die Ermittlung des Goldstandards in 90 Fällen durch Laparotomie (davon 64 mit Resektion, die restlichen 26 explorativ mit Biopsieentnahme) und in 16 Fällen durch anderweitige Gewebeentnahme. In den restlichen 13 Fällen erfolgte die Diagnose indirekt

Diagnose	
Adenokarzinom	83 (%)
anderes Karzinom	21 (%)
Karzinom ohne Histologie	15 (%)
chronische Pankreatitis	55 (%)
Zysten	9 (%)
andere benigne Erkrankung	19 (%)
UICC-Klassifikation	
I	7 (%)
II	10 (%)
III	48 (%)
IVa	18 (%)
IVb	36 (%)

Tabelle 3.1: *Diagnosen der 193 Patienten, die zur Analyse herangezogen wurden. Angegeben sind absolute und relative Häufigkeiten.*

über den klinischen Verlauf: Die Patienten hatten Leber- oder andere Metastasen und starben innerhalb von sechs Monaten an Tumorkachexie oder anderen tumorbedingten Komplikationen.

Bei 47 der 64 resezierten Patienten mit maligner Läsion konnte der Tumor komplett entfernt werden (R0). Bei 16 Patienten verblieben mikroskopische Residualtumore an den Schnittträndern (R1) und bei einem Patienten konnten nur Teile des Tumors entfernt werden (R2). Dieser Tumor wurde als irresektabel eingestuft. Von den 25 nicht resezierten Patienten hatten elf

Lokalisation	
Pankreaskopf	165 (85%)
Körper	13 (7%)
Schwanz	6 (3%)
keine Angabe	8 (%)
Größe	
< 2cm	29 (%)
2-3cm	48 (%)
3-4cm	51 (%)
< 4cm	35 (%)
keine Angabe	30 (%)

Tabelle 3.2: Lokalisation und Größe der 193 Läsionen, die zur Analyse herangezogen wurden. Angegeben sind absolute und relative Häufigkeiten.

Fernmetastasen, zehn waren lokal irresektabel und drei Patienten hatten sowohl Fernmetastasen als auch einen lokal irresektablen Tumor. Da bei den elf Patienten mit Fernmetastasen die Operation abgebrochen wurde, liegen keine Angaben zur lokalen Resektabilität vor. Die zwei restlichen Patienten konnten wegen während der Operation festgestellter Zirrhose bzw. portaler Hypertension nicht reseziert werden. Somit lag nur bei 77 Patienten (47 R0, 16 R1, 1 R2, 10 lokal irresektabel und drei lokal irresektabel mit Fernmetastasen) der Goldstandard für die lokale Resektabilität vor.

Bei den 74 Patienten mit benigner Läsion wurde in 49 Fällen eine Laparotomie durchgeführt, in fünf Fällen lag eine Biopsie vor. Bei 16 Patienten ergaben Kontrolluntersuchungen keine Veränderung der Läsion über einen

Zeitraum von mindestens einem Jahr. Vier Patienten lehnten Kontrolluntersuchungen ab, konnten aber mindestens zwei Jahre nachbeobachtet werden, so dass auch hier ein malignes Pankreaskarzinom ausgeschlossen werden konnte.

### 3.1 Vergleich der Modalitäten

Nicht alle diagnostischen Tests konnten bei allen Patienten protokollgemäß durchgeführt werden. Bei einigen Patienten lagen bereits Ergebnisse externer Tests vor und die Patienten weigerten sich, die Untersuchung unter Studienbedingungen zu wiederholen. Auch aus organisatorischen Gründen (hauptsächlich Gerätedefekte) konnten manche Untersuchungen nicht durchgeführt werden. Dies kam bei EUS 32 Mal vor, bei ERCP 18 Mal, zwölf Mal bei MR, elf Mal bei CT, vier Mal bei PET und zwei Mal bei US. Um einen möglichen Selektionsbias auszuschließen, wurden die Analysen sowohl für alle 193 Patienten als auch für die Gruppe der 133 Patienten durchgeführt, die alle sechs Untersuchungen komplett absolvierten.

Des Weiteren gab es bei allen Modalitäten Fälle, in denen der Test methodenbedingt nicht durchgeführt werden konnte oder die Ergebnisse nicht auswertbar waren. Hauptsächlich trat dieses Problem bei US auf, da bei 19 Untersuchungen (10%) die Bauchspeicheldrüse aufgrund von Gasüberlagerungen nicht darstellbar war. Bei CT erfolgten vier Untersuchungen ohne Kontrastmittel, da bei den Patienten eine Kontrastmittelallergie vorlag. MR konnte je einmal wegen Metallimplantaten bzw. Klaustrophobie nicht durchgeführt werden. Bei ERCP konnte die Papille in einem Fall wegen einer Billroth-II-Anatomie nicht erreicht werden. EUS konnte aus anatomischen

Gründen zweimal nicht durchgeführt werden, bei einem Patienten war die PET-Untersuchung nicht auswertbar. Dementsprechend basieren die Auswertungen auf zwei Populationen, der intention-to-diagnose-Population (vergleichbar mit der intention-to-treat-Population bei Therapiestudien) und der per-protocol-Population.

Signifikante Komplikationen traten in zwei Fällen bei ERCP auf, alle anderen Untersuchungen verliefen komplikationsfrei.

### **3.1.1 Erkennen des Pankreaskarzinoms**

Die diagnostische Leistungsfähigkeit der sechs Einzelverfahren bei der intention-to-diagnose-(itd)-Analyse in der Gruppe der 193 Patienten ist in Tabelle 3.3 dargestellt. ERCP und MR haben mit 95% bzw. 91% die höchste Sensitivität und sind statistisch signifikant besser als US ( $p < 0,001$  bei beiden), EUS ( $p < 0,001$  bzw.  $p = 0,001$ ) und CT ( $p = 0,004$  bzw.  $0,031$ ). Die Sensitivität von ERCP ist zudem signifikant besser als PET ( $p = 0,017$ ). Der US hat mit 66% die niedrigste Sensitivität, hat aber in der per-protocol-(pp)-Analyse, d.h. nach Ausschluss der Fälle, in denen das Pankreas nicht dargestellt werden konnte, eine höhere Sensitivität von 73% und ist dann nicht mehr signifikant schlechter als EUS und CT.

EUS hat mit 82% die höchste Spezifität, US mit 64% die niedrigste. Bezüglich der Spezifität ist nur der Unterschied zwischen US und EUS statistisch signifikant. In der pp-Analyse steigt die Spezifität des US auf 72% und ist dann nicht mehr signifikant schlechter als EUS.

Wenn nur die 133 Patienten, bei denen alle Tests durchgeführt wurden, be-

	Sens	Spez
US	66% (57%-74%)	64% (53%-74%)
EUS	72% (62%-80%)	82% (72%-90%) <sup>a</sup>
ERCP	95% (89%-98%) <sup>abcd</sup>	71% (59%-81%)
CT	81% (72%-87%) <sup>a</sup>	78% (67%-86%)
MR	91% (84%-95%) <sup>abc</sup>	76% (65%-85%)
PET	83% (75%-89%) <sup>a</sup>	69% (57%-78%)

Tabelle 3.3: *Sensitivität und Spezifität mit dazugehörigen 95%-Konfidenzintervallen der sechs Modalitäten bezüglich der Dignität bei der itd-Analyse.* <sup>a</sup> *stat. signifikant besser als US,* <sup>b</sup> *stat. signifikant besser als EUS,* <sup>c</sup> *stat. signifikant besser als CT,* <sup>d</sup> *stat. signifikant besser als PET.*

trachtet werden, unterscheiden sich die Ergebnisse nur geringfügig: Die Sensitivität bzw. Spezifität beträgt bei US 63% bzw. 69% (70% bzw. 78% bei pp), bei EUS 73% bzw. 80%, bei ERCP 95% bzw. 73%, bei CT 81% bzw. 78%, bei MR 92% bzw. 77% und bei PET 83% bzw. 69%. Die Konfidenzintervalle sind durch die kleineren Fallzahlen etwas breiter.

ERCP und MR haben ein negatives Likelihood Ratio kleiner 0,1 (siehe Tabelle 3.4), d.h. negative Ergebnisse haben eine große klinische Relevanz (vgl. Tabelle 2.2). Alle anderen Modalitäten liegen zwischen 0,2 und 0,5, d.h. negative Resultate bei diesen Modalitäten sind nur mäßig aussagekräftig. Da die positiven Likelihood Ratios von MR und CT größer sind als fünf, haben positive Testergebnisse bei beiden Modalitäten eine hohe Aussagekraft. Bei den anderen Modalitäten mit Werten zwischen zwei und fünf haben positive Testergebnisse hingegen nur eine mäßige Aussagekraft.

	LR <sup>+</sup>	LR <sup>-</sup>
US	3,71 (2,23-6,18)	0,33 (0,22-0,48)
EUS	4,45 (2,55-7,76)	0,31 (0,22-0,45)
ERCP	3,67 (2,43-5,54)	0,08 (0,03-0,17)
CT	6,17 (3,33-11,43)	0,20 (0,13-0,32)
MR	5,45 (3,24-9,17)	0,09 (0,05-0,19)
PET	2,63 (1,85-3,72)	0,24 (0,15-0,37)

Tabelle 3.4: *Likelihood Ratios mit dazugehörigen 95%-Konfidenzintervallen der sechs Modalitäten bezüglich der Dignität bei der itd-Analyse.*

### 3.1.2 Erkennen der Resektabilität

Die Ergebnisse der vier Modalitäten bei der Beurteilung der lokalen Resektabilität sind in Tabelle 3.5 dargestellt. Die Sensitivität, d.h. das Erkennen eines resektablen Tumors war für MR und EUS größer als 90% und lag für CT und US bei 83% bzw. 84%. In der pp-Population erhöht sich die Sensitivität des US auf 95%, bleibt aber für CT mit 84% unverändert. CT klassifiziert sieben der 58 resektablen Tumore fälschlicherweise als irresektabel (13%). Diese Rate war bei den anderen Modalitäten deutlich geringer (US: 3 von 62 (5%), EUS: 1 von 50 (2%), MR: 1 von 60 (12%)). Der Unterschied zwischen MR und CT bzw. US ist statistisch signifikant ( $p=0,012$  bzw.  $p=0,008$ ).

Die Spezifität, d.h. das Erkennen der Irresektabilität war bei allen vier Modalitäten niedrig. CT liegt bei 57%, alle anderen Verfahren unter 50%. Die Unterschiede zwischen den Modalitäten sind statistisch nicht signifikant.

In Tabelle 3.8 sind die Likelihood Ratios der sechs Modalitäten dargestellt.

	Sens	Spez
US	84% (57%-74%)	29% (11%-55%)
EUS	96% (62%-80%)	20% (6%-51%)
CT	83% (72%-87%)	57% (33%-79%)
MR	98% (84%-95%) <sup>ab</sup>	40% (17%-69%)

Tabelle 3.5: *Diagnostische Leistungsfähigkeit der sechs Einzelverfahren bezüglich der lokalen Resektabilität bei der itd-Analyse. Angegeben sind Sensitivität und Spezifität mit dazugehörigen Konfidenzintervallen. <sup>a</sup> stat. signifikant besser als US, <sup>b</sup> stat. signifikant besser als CT*

	LR <sup>+</sup>	LR <sup>-</sup>
US	3,71 (2,23-6,18)	0,33 (0,22-0,48)
EUS	4,45 (2,55-7,76)	0,31 (0,22-0,45)
CT	6,17 (3,33-11,43)	0,20 (0,13-0,32)
MR	5,45 (3,24-9,17)	0,09 (0,05-0,19)

Tabelle 3.6: *Diagnostische Leistungsfähigkeit der sechs Einzelverfahren bezüglich der lokalen Resektabilität bei der itd-Analyse. Angegeben sind die Likelihood Ratios mit dazugehörigen Konfidenzintervallen.*

Das positive Likelihood Ratio ist für CT und MR größer als fünf, d.h. bei beiden Modalitäten erhöht ein positives Ergebnis die Evidenz deutlich, bei US und EUS mit Werten kleiner fünf hingegen nur mäßig. Da die negativen Likelihood Ratios für US, EUS und CT größergleich 0,2 sind, tragen negative Ergebnisse dieser Modalitäten nur mäßig zum Erkenntnisgewinn bei. Negative Ergebnisse bei MR haben aufgrund eines negativen Likelihood Ratio Wertes kleiner 0,1 hingegen eine hohe Relevanz.

Patienten mit bereits diagnostizierten Fernmetastasen wurden in die Studie nicht eingeschlossen. Allerdings wurden bei 29 Patienten im Laufe der Studie Fernmetastasen detektiert: bei 15 Patienten während der präoperativen Diagnostik (alle in der Leber) und bei 14 Patienten während der chirurgischen Exploration (in sieben Fällen in der Leber, in fünf Fällen im Peritoneum und in zwei Fällen sowohl in der Leber als auch im Peritoneum). Alle diagnostisch nicht entdeckten Fernmetastasen waren kleiner als ein Zentimeter. Die Sensitivitäten der sechs Modalitäten für die Entdeckung von Fernmetastasen sind in Tabelle (3.7) dargestellt. Die Unterschiede zwischen den Modalitäten sind alle statistisch nicht signifikant.

	Sensitivität (KI)
US	38% (23% - 56%)
EUS	25% (11% - 47%)
CT	43% (27% - 61%)
MR	36% (21% - 54%)
PET	38% (23% - 56%)

Tabelle 3.7: *Diagnostische Leistungsfähigkeit der sechs Einzelverfahren bezüglich der Detektion von Fernmetastasen bei der itd-Analyse. Angegeben sind die Sensitivitäten mit dazugehörigen Konfidenzintervallen.*

## 3.2 Kombination der Modalitäten

Da bildgebende Verfahren prinzipiell nur nacheinander durchgeführt werden, ist eine Zeitersparnis nicht möglich. Zudem ist es wünschenswert, nur so viele Patienten wie notwendig mehrfach zu testen, da die Methoden zum Teil recht

invasiv und teuer sind. Somit wird im vorliegenden Fall sequentiell getestet.

Da die Sensitivität bei der Bestimmung der Dignität bei den besten Modalitäten über 90% liegt, die Spezifität hingegen bei allen Verfahren kleiner als 80% ist, soll durch die Kombination von Modalitäten die Spezifität verbessert werden. Dies geschieht durch Verwenden der BTN-Regel (vgl. Seite 35).

Die „naive“ Kombination des Verfahrens mit der größten Spezifität (EUS mit 82%) mit dem Verfahren mit der höchsten Sensitivität (ERCP mit 95%) mittels BTN-Regel führt zu einer erhöhten Spezifität von 85%, auf Kosten einer deutlich verringerten Sensitivität von 53%.

Da durch Verwendung der BTN-Regel nach (2.16) zwar die Spezifität verbessert wird, die Sensitivität der Kombination aber unter dem Minimum der beiden Einzelsensitivitäten liegt, werden bei der empfohlenen Kombination (38) die Tests mit den höchsten Sensitivitäten (ERCP mit 95% und MR mit 91%) kombiniert, um den Sensitivitätsverlust möglichst gering zu halten. Die Sensitivität der Kombination beträgt 72% und liegt damit über der Sensitivität von US und gleichauf mit EUS. Die Spezifität ist nach (2.16) größer als das Maximum der beiden Einzelspezifitäten (MR: 76% und ERCP: 71%) und beträgt 91%.

Bei der Beurteilung der Resektabilität soll durch die Kombination ebenfalls die Spezifität, d.h. das Erkennen der Irresektabilität, verbessert werden. Da alle vier Einzelsensitivitäten über 80% liegen, ist bei allen möglichen Kombinationen - abgesehen von US kombiniert mit EUS - der Sensitivitätsverlust durch die Kombination mittels BTN-Regel akzeptabel, der Spezifitätsgewinn

jedoch deutlich (siehe Tabelle 3.8).

	Sens	Spez
US + EUS	42% (19%-68%)	76% (63%-86%)
US + CT	71% (45%-88%)	68% (59%-79%)
US + MR	73% (43%-90%)	82% (70%-89%)
EUS + CT	73% (43%-90%)	76% (63%-86%)
EUS + MR	63% (31%-86%)	94% (83%-98%)
CT + MR	67% (39%-86%)	82% (70%-90%)

Tabelle 3.8: *Beurteilung der lokalen Resektabilität bei Kombination mittels BTN-Regel. Angegeben sind Sensitivität und Spezifität mit dazugehörigen Konfidenzintervallen.*

Die Kombination von US und MR besitzt sowohl eine höhere Sensitivität als auch eine höhere Spezifität als vier der fünf Kombinationen. Die Kombination EUS und MR erreicht als „optimale“ Kombination mit 94% erwartungsgemäß die höchste Sensitivität, weist aber mit 63% eine deutlich geringere Spezifität auf.

Nur wenn die Korrelation der beiden kombinierten Tests innerhalb der Kranken hoch und innerhalb der Gesunden niedrig ist, kann durch Kombination mittels BTN-Regel sowohl der positive wie auch der negative prädiktive Wert verbessert werden (vgl. Seite 43). Da sich die Korrelationen von je zwei Tests innerhalb der Subpopulation der Malignen kaum von den Korrelationen innerhalb der Subpopulation der Benigen unterscheiden (Tabelle 3.9), haben alle Kombinationen mittels BTN-Regel einen niedrigeren negativen und einem höheren positiven prädiktiven Wert als mindestens einer der beiden

Einzeltests.

	US	EUS	CT	MR	ERCP
PET	0,047/0,217	0,067/0,147	0,159/0,269	0,315/0,255	0,003/0,186
US		0,377/0,523	0,338/0,288	0,354/0,241	0,006/0,221
EUS			0,161/0,505	0,254/0,453	0,235/0,490
CT				0,387/0,420	0,205/0,405
MR					0,082/0,346

Tabelle 3.9:  $\phi$ -Koeffizienten als Maß der Korrelation der Tests innerhalb der Subpopulation der Malignen bzw. Benignen

Bei den meisten Kombinationen kann man, wie in Tabelle 3.10 dargestellt, von einer bedingten Unabhängigkeit beider Einzeltests ausgehen. Dementsprechend verschlechtert sich der positive prädiktive Wert durch die Kombination mittels BTN-Regel nicht. Aber auch bei Tests, die, wie z.B. CT und MR, bedingt abhängig sind, ist die Abhängigkeit nicht so groß, dass der positive prädiktive Wert abnimmt.

### 3.3 Klinische Entscheidungsanalyse

#### 3.3.1 Basisfallanalyse

Die optimale Kombinationsregel (BTN oder BTP) hing von den zu kombinierenden Modalitäten ab. Aber alle homogenen Strategien (BTN oder BTP sowohl für die Detektion des Pankreaskarzinoms als auch für die Beurteilung der Resektabilität), erzielten bessere Resultate als gemischte Regeln (BTN

	US	EUS	CT	MR	ERCP
PET	1,000/0,794	1,000/0,171	0,207/0,204	0,035/0,033	1,000/0,771
US		0,007/0,049	0,008/0,529	0,326/0,553	1,000/0,154
EUS			0,385/0,001	0,079/0,056	0,071/0,001
CT				0,003/0,010	0,325/0,047
MR					1,000/0,236

Tabelle 3.10: *p*-Werte des exakten Tests nach Fisher innerhalb der Subpopulation der Malignen bzw. Benigen

für die Detektion des Pankreaskarzinoms und BTP für die Beurteilung der Resektabilität oder umgekehrt). Für den Vergleich der verschiedenen Kombinationen wurde für jede Kombination die beste Kombinationsregel ausgewählt. Somit reduziert sich die Zahl der Strategien von 44 auf 18 (siehe Seite 74).

Alle sechs Strategien mit MR als erster Modalität waren unter den besten acht Strategien (3.1). Die beste diagnostische Leistung erbrachte die Strategie „MR gefolgt von ERCP für positive Ergebnisse“, die 78% der Patienten richtig in benigne, maligne/resektabel und maligne/irresektabel klassifizierte. Die zweitbeste Strategie war „nur MR“, mit 77% richtig klassifizierten Patienten. Die Kombination „US gefolgt von PET für negative Ergebnisse“ lieferte mit nur 61% korrekt klassifizierten Patienten die schlechtesten Ergebnisse.

Die Prozentzahlen korrekt klassifizierter Patienten unterscheiden sich nicht deutlich zwischen den Strategien. Allerdings gibt es beträchtliche Unterschie-

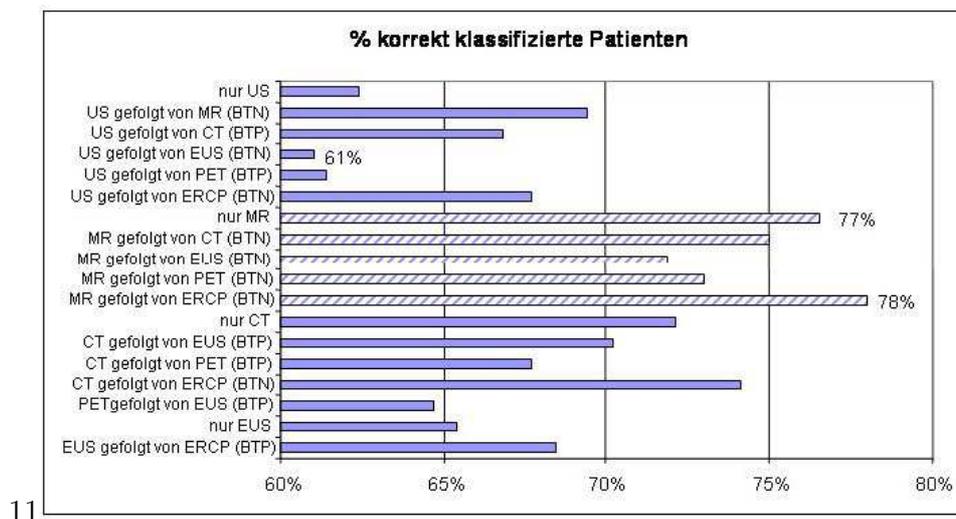


Abbildung 3.1: Diagnostische Leistungsfähigkeit der 18 Strategien. Dargestellt sind die Anteile richtig klassifizierter Patienten in %

de zwischen den Strategien bezüglich der verschiedenen Arten von Fehlklassifikationen. In Abbildung (3.2) sind die Konsequenzen der besten Strategie „MR gefolgt von ERCP für positive Ergebnisse“ im Vergleich zur zweitbesten Strategie „nur MR“ und der besten Strategie mit CT als erster Modalität „CT gefolgt von ERCP für positive Ergebnisse“ dargestellt.

Bei einer hypothetischen Kohorte von 1000 Patienten und der Prävalenz der Studie würden bei Verwendung der Strategie „MR gefolgt von ERCP für positive Ergebnisse“ 128 Patienten unnötigerweise operiert (davon 101 Patienten mit in Wirklichkeit irresektablem Tumor und 27 Patienten mit benigner Erkrankung) und 51 Patienten mit resektablem Tumor wurden fälschlicherweise nicht operiert (davon 14 fälschlicherweise als irresektabel und 37 als benigne klassifiziert).

Wenn positive Ergebnisse nicht mit ERCP verifiziert werden (Strategie „nur

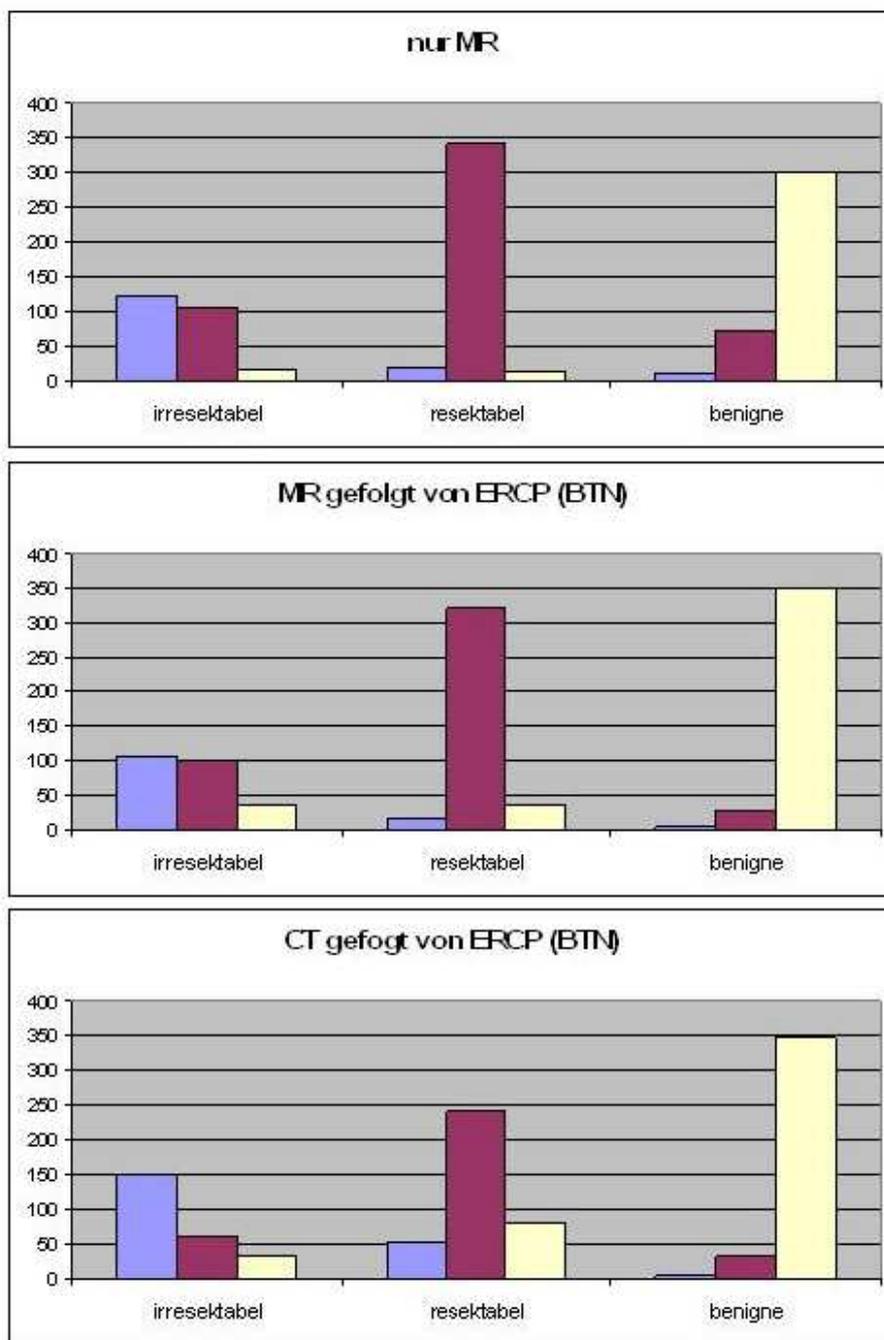


Abbildung 3.2: Darstellung der Konsequenzen der drei Strategien „MR gefolgt von ERCP für positive Ergebnisse“, „nur MR“ und „CT gefolgt von ERCP für positive Ergebnisse“ in einer hypothetischen Kohorte von 1.000 Patienten. Die Balken repräsentieren die Anzahl der Patienten, die Chemotherapie erhalten (blau), reseziert werden (rot) oder anderweitig behandelt werden (gelb).

MR“) würden weitere 51 Patienten unnötigerweise operiert (davon sieben mit eigentlich irresektablem Karzinom und 44 mit benigner Erkrankung). Allerdings würden 20 zusätzliche Patienten mit resektablem Tumor richtigerweise operiert.

Mit der Strategie „CT gefolgt von ERCP für positive Ergebnisse“, würden im Vergleich zur Strategie „nur MR“ 80 Patienten die Operation vorenthalten, während 36 Patienten eine unnötige Operation erspart bleiben würde.

### **3.3.2 Sensitivitätsanalysen**

Einweg-Sensitivitätsanalysen wurden bezüglich der Prävalenz durchgeführt. Wenn die Prävalenz wie in der Studie und der Basisfallanalyse geringer ist als 69%, hat die Strategie „MR gefolgt von ERCP“ für negative Ergebnisse den höchsten erwarteten Wert korrekt klassifizierter Patienten. Wenn die Prävalenz größer ist als 69%, ist die Strategie „nur MR“ allen anderen Strategien überlegen.

Die Ergebnisse waren robust gegenüber simultanen Variationen der Modellparameter. In einer Monte Carlo Simulation mit 100.000 Stichproben war „MR gefolgt von ERCP für positive Ergebnisse“ in 54% der Fälle die optimale Strategie (siehe Abbildung 3.3). Im Mittel wurden 78% der Patienten richtig zugeordnet (mit einer Standardabweichung von 2,9%). Alle anderen Strategien waren in weniger als 15% der Fälle optimal.

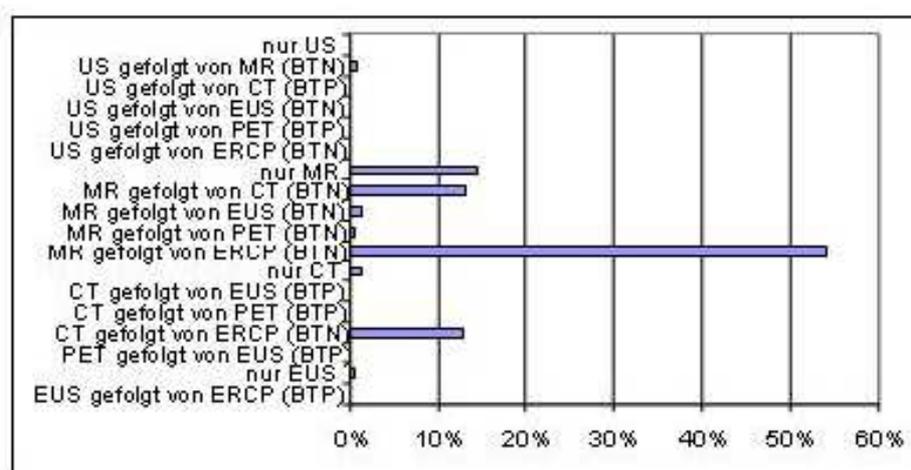
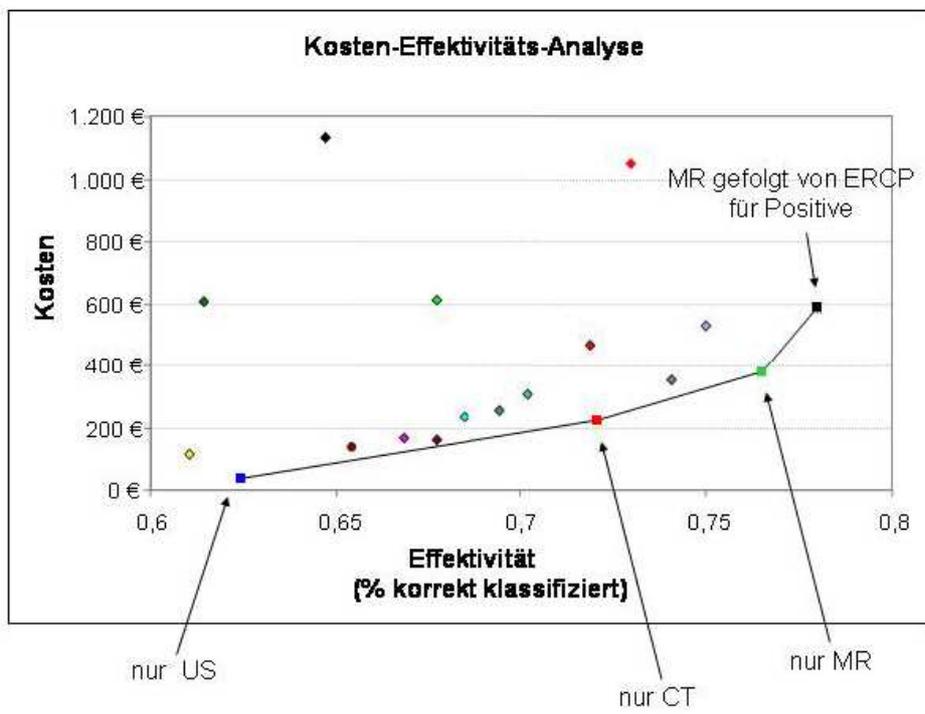


Abbildung 3.3: Ergebnisse der probabilistischen Sensitivitätsanalyse. Dargestellt sind die relativen Häufigkeiten, mit denen die einzelnen Strategien bei den 100.000 Simulation optimal waren.

### 3.3.3 Kosten-Effektivitäts-Analyse

Wenn die Kosten der einzelnen diagnostischen Verfahren berücksichtigt werden, ergibt sich ein ähnliches Bild. Alle Strategien außer „nur US“, „nur CT“, „nur MR“ und „MR gefolgt von ERCP für Positive“ werden dominiert (siehe Abbildung 3.4).

In Tabelle 3.11 sind die Ergebnisse der Kosten-Effektivitäts-Analyse für die vier nicht dominierten Strategien dargestellt. Die Strategie „nur US“ ist mit erwarteten Kosten von 37,90 Euro am billigsten, pro richtig diagnostiziertem Patienten entstehen Kosten von 61 Euro. Die erwarteten Kosten der Strategie „nur CT“ liegen um 185,40 Euro höher, es werden aber auch 9,7% mehr Patienten korrekt diagnostiziert. Dies führt zu Kosten pro richtig diagnostiziertem Patienten von 310 Euro. Die inkrementellen erwarteten Kosten von CT bezogen auf US zu den korrespondierenden inkrementellen medizinischen



11

Abbildung 3.4: Kosteneffektivität der 18 Strategien

Effekten (ICER) betragen 1.911 Euro. Die Strategie „nur MR“ ist um weitere 154,70 Euro teurer als die Strategie „nur CT“, klassifiziert aber 4,4% mehr Patienten richtig. Die Kosten für jeden im Vergleich zu CT zusätzlich richtig klassifizierte Patienten (ICER) belaufen sich auf 3.516 Euro. Wenn positive Ergebnisse des MR mit ERCP verifiziert werden (Strategie „MR gefolgt von ERCP für Positive“), erhöht sich der Anteil der richtig klassifizierten Patienten um weitere 1,5%, allerdings steigen die erwarteten Kosten um beträchtliche 587,10 Euro. Die Kosten für jeden im Vergleich zur Strategie „nur MR“ zusätzlich richtig diagnostizierten Patienten betragen 13.490 Euro.

Strategie	K	$\Delta K$	E	$\Delta E$	C/E	ICER
nur US	37,90		0,62		61	
nur CT	223,30	185,40	0,72	0,097	310	1.911
nur MR	378,00	154,70	0,77	0,044	494	3.516
MR und ERCP	587,10	209,10	0,78	0,015	753	13.940

Tabelle 3.11: *Ergebnisse der Kosten-Effektivitätsanalyse für die vier nicht dominierten Strategien. Dargestellt sind Kosten  $K$ , inkrementelle Kosten  $\Delta K$ , Effektivität  $E$ , inkrementelle Effektivität  $\Delta E$ , das Kosten-Effektivitätsverhältnis  $C/E$  und das inkrementelle Kosten-Effektivitätsverhältnis ICER.*