Mining Oncology Data: Knowledge Discovery in Clinical Performance of Cancer Patients

by

John Hayward

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Master of Science

 in

Computer Science

August 2006

APPROVED:

Professor Carolina Ruiz, Thesis Advisor

Professor Sergio Alvarez (Boston College), Thesis Co-Advisor

Professor George Heineman, Thesis Co-Advisor

Dr. Giles Whalen (U. Massachusetts Medical School), Thesis Co-Advisor

Professor Murali Mani, Thesis Reader

Professor Michael A. Gennert, Head of Department

Abstract

Our goal in this research is twofold: to develop clinical performance databases of cancer patients, and to conduct data mining and machine learning studies on collected patient records. We use these studies to develop models for predicting cancer patient medical outcomes. The clinical database is developed in conjunction with surgeons and oncologists at UMass Memorial Hospital. Aspects of the database design and representation of patient narrative are discussed here. Current predictive model design in medical literature is dominated by linear and logistic regression techniques. We seek to show that novel machine learning methods can perform as well or better than these traditional techniques.

Our machine learning focus for this thesis is on pancreatic cancer patients. Classification and regression prediction targets include patient survival, wellbeing scores, and disease characteristics. Information research in oncology is often constrained by type variation, missing attributes, high dimensionality, skewed class distribution, and small data sets. We compensate for these difficulties using preprocessing, meta-learning, and other algorithmic methods during data analysis. The predictive accuracy and regression error of various machine learning models are presented as results, as are t-tests comparing these to the accuracy of traditional regression methods. In most cases, it is shown that the novel machine learning prediction methods offer comparable or superior performance. We conclude with an analysis of results and discussion of future research possibilities.

Contents

1	Intr	oducti	ion	1
2	Med	dical B	Background	3
	2.1	Pancre	eatic Cancer Background	7
3	Clin	nical D	atabase Construction	15
	3.1	Gastro	pintestinal Cancer Database	15
		3.1.1	Pancreatic Cancer	17
		3.1.2	Hepatocellular Cancer	31
		3.1.3	Gall Bladder/Biliary Cancer	45
		3.1.4	Gastric Cancer	58
		3.1.5	Esophageal Cancer	71
		3.1.6	Colorectal Cancer	85
	3.2	Breast	Cancer Database	99
4	Clin	nical P	erformance Machine Learning -	
	Pro	cedure	e & Design	112
	4.1	Object	tives of Analysis	112
	4.2	Patien	nt Data Set	113
	4.3	Data I	Mining and Machine Learning Algorithms Used	118
		4.3.1	Benchmark Algorithms	118
		4.3.2	Classification Algorithms	119
		4.3.3	Regression Algorithms	120
		4.3.4	Data Preprocessing Algorithms	121
		4.3.5	Meta-Learning Algorithm	122
	4.4	Exper	imental Design	124
		4.4.1	Classification Tests	126

		4.4.2 Regression Tests	128
5	Clin	ical Performance Machine Learning -	
	Res	ults & Analysis	130
	5.1	C1 - Tumor Size	130
	5.2	C2 - T-Stage	136
	5.3	C3 - N-Stage	141
	5.4	C4 - Vascular Involvement	146
	5.5	C5 - Histology	151
	5.6	C6 - Malignancy	156
	5.7	C7 - ECOG 6-Month	161
	5.8	C8 - ECOG 9-Month	166
	5.9	C9 - ECOG 12-Month	171
	5.10	C10 - Survival	176
	5.11	R1 - Tumor Size	181
	5.12	R2 - ECOG 6-Month	185
	5.13	R3 - ECOG 9-Month	189
	5.14	R4 - ECOG 12-Month	193
	5.15	R5 - Survival	197
6	Hig	h-Performance Predictive Models	201
	6.1	Classification - Histology - Data Set C - Bayesian Net 2-Parent	201
	6.2	Classification - Survival - Data Set C - Bayesian Net 2-Parent	214
	6.3	Regression - ECOG 6-Month - Data Set F - Linear Regression w/ Bagging $% \mathcal{A} = \mathcal{A} = \mathcal{A}$.	223
	6.4	Regression - ECOG 9-Month - Data Set F - Multi-layer Perceptron w/ 2	
		Hidden Layers	233
7	Rela	ated Work	240

List of Figures

1	Gray's Anatomy - Pancreas and Periampullary Region [Gra95]	8
2	Digestive System with Common Bile Duct Illustrated [Cen]	9
3	Whipple Procedure - Pre-Surgical Anatomy [Cli]	13
4	Whipple Procedure - Post-Surgical Anatomy [Cli]	14
5	Pancreatic Cancer Presentation Form	17
6	Pancreatic Cancer Presentation Schema	18
7	Pancreatic Cancer Medical History Form	18
8	Pancreatic Cancer Medical History Table Schema	19
9	Pancreatic Cancer Serums Studies Form	20
10	Pancreatic Cancer Serums Studies Table Schema	20
11	Pancreatic Cancer Diagnostic Imaging Form	21
12	Pancreatic Cancer Diagnostic Imaging Table Schema	21
13	Pancreatic Cancer Endoscopy Studies Form	22
14	Pancreatic Cancer Endoscopy Studies Table Schema	23
15	Pancreatic Cancer Preliminary Outlook Form	23
16	Pancreatic Cancer Preliminary Outlook Table Schema	23
17	Pancreatic Cancer Treatment Form	24
18	Pancreatic Cancer Treatment Table Schema	25
19	Pancreatic Cancer Resection Form	26
20	Pancreatic Cancer Resection Table Schema	26
21	Pancreatic Cancer No Resection Form	27
22	Pancreatic Cancer No Resection Table Schema	27

 $\mathbf{244}$

23	Pancreatic Cancer Pathology Form	28
24	Pancreatic Cancer Pathology Table Schema	28
25	Pancreatic Cancer Follow-Up Form	29
26	Pancreatic Cancer Follow-Up Table Schema	30
27	Hepatocellular Cancer Presentation Form	31
28	Hepatocellular Cancer Presentation Table Schema	32
29	Hepatocellular Cancer Medical History Form	32
30	Hepatocellular Cancer Medical History Table Schema	33
31	Hepatocellular Cancer Serum Studies Form	34
32	Hepatocellular Cancer Serum Studies Table Schema	34
33	Hepatocellular Cancer Diagnostic Imaging Form	35
34	Hepatocellular Cancer Diagnostic Imaging Table Schema	35
35	Hepatocellular Cancer Preliminary Outlook Form	36
36	Hepatocellular Cancer Preliminary Outlook Table Schema	36
37	Hepatocellular Cancer Treatment Form	37
38	Hepatocellular Cancer Treatment Table Schema	37
39	Hepatocellular Cancer Ablation Form	38
40	Hepatocellular Cancer Ablation Table Schema	38
41	Hepatocellular Cancer Resection Form	39
42	Hepatocellular Cancer Resection Table Schema	40
43	Hepatocellular Cancer No Resection Form	41
44	Hepatocellular Cancer No Resection Table Schema	41
45	Hepatocellular Cancer Pathology Form	42
46	Hepatocellular Cancer Pathology Table Schema	42
47	Hepatocellular Cancer Follow-Up Form	43
48	Hepatocellular Cancer Follow-Up Table Schema	44
49	Gall Bladder/Biliary Cancer Presentation Form	45

50	Gall Bladder/Biliary Cancer Presentation Table Schema	46
51	Gall Bladder/Biliary Cancer Medical History Form	46
52	Gall Bladder/Biliary Cancer Medical History Table Schema	47
53	Gall Bladder/Biliary Cancer Serum Studies Form	48
54	Gall Bladder/Biliary Cancer Serum Studies Table Schema	48
55	Gall Bladder/Biliary Cancer Diagnostic Imaging Form	49
56	Gall Bladder/Biliary Cancer Diagnostic Imaging Table Schema	49
57	Gall Bladder/Biliary Cancer Preliminary Outlook Form	50
58	Gall Bladder/Biliary Cancer Preliminary Outlook Table Schema	50
59	Gall Bladder/Biliary Cancer Treatment Form	51
60	Gall Bladder/Biliary Cancer Treatment Table Schema	51
61	Gall Bladder/Biliary Cancer Resection Form	52
62	Gall Bladder/Biliary Cancer Resection Table Schema	53
63	Gall Bladder/Biliary Cancer No Resection Form	54
64	Gall Bladder/Biliary Cancer No Resection Table Schema	54
65	Gall Bladder/Biliary Cancer Pathology Form	55
66	Gall Bladder/Biliary Cancer Pathology Table Schema	55
67	Gall Bladder/Biliary Cancer Follow-Up Form	56
68	Gall Bladder/Biliary Cancer Follow-Up Table Schema	57
69	Gastric Cancer Presentation Form	58
70	Gastric Cancer Presentation Table Schema	59
71	Gastric Cancer Medical History Form	59
72	Gastric Cancer Medical History Table Schema	60
73	Gastric Cancer Serum Studies Form	61
74	Gastric Cancer Serum Studies Table Schema	61
75	Gastric Cancer Diagnostic Imaging Form	62
76	Gastric Cancer Diagnostic Imaging Table Schema	62

77	Gastric Cancer Preliminary Outlook Form	63
78	Gastric Cancer Preliminary Outlook Table Schema	63
79	Gastric Cancer Treatment Form	64
80	Gastric Cancer Treatment Table Schema	64
81	Gastric Cancer Resection Form	65
82	Gastric Cancer Resection Table Schema	66
83	Gastric Cancer No Resection Form	67
84	Gastric Cancer No Resection Table Schema	67
85	Gastric Cancer Pathology Form	68
86	Gastric Cancer Pathology Table Schema	68
87	Gastric Cancer Follow-Up Form	69
88	Gastric Cancer Follow-Up Table Schema	70
89	Esophageal Cancer Presentation Form	71
90	Esophageal Cancer Presentation Table Schema	72
91	Esophageal Cancer Medical History Form	72
92	Esophageal Cancer Medical History Table Schema	73
93	Esophageal Cancer Serum Studies Form	74
94	Esophageal Cancer Serum Studies Table Schema	74
95	Esophageal Cancer Diagnostic Imaging Form	75
96	Esophageal Cancer Diagnostic Imaging Table Schema	75
97	Esophageal Cancer Preliminary Outlook Form	76
98	Esophageal Cancer Preliminary Outlook Table Schema	76
99	Esophageal Cancer Treatment Form	77
100	Esophageal Cancer Treatment Table Schema	78
101	Esophageal Cancer Resection Form	79
102	Esophageal Cancer Resection Table Schema	80
103	Esophageal Cancer No Resection Form	81

104	Esophageal Cancer No Resection Table Schema	81
105	Esophageal Cancer Pathology Form	82
106	Esophageal Cancer Pathology Table Schema	82
107	Esophageal Cancer Follow-Up Form	83
108	Esophageal Cancer Follow-Up Table Schema	84
109	Colorectal Cancer Presentation Form	85
110	Colorectal Cancer Presentation Table Schema	86
111	Colorectal Cancer Medical History Form	86
112	Colorectal Cancer Medical History Table Schema	87
113	Colorectal Cancer Serum Studies Form	88
114	Colorectal Cancer Serum Studies Table Schema	88
115	Colorectal Cancer Diagnostic Imaging Form	89
116	Colorectal Cancer Diagnostic Imaging Table Schema	89
117	Colorectal Cancer Preliminary Outlook Form	90
118	Colorectal Cancer Preliminary Outlook Table Schema	91
119	Colorectal Cancer Treatment Form	91
120	Colorectal Cancer Treatment Table Schema	92
121	Colorectal Cancer Ablation Form	93
122	Colorectal Cancer Ablation Table Schema	93
123	Colorectal Cancer Resection Form	94
124	Colorectal Cancer Resection Table Schema	95
125	Colorectal Cancer No Resection Form	95
126	Colorectal Cancer No Resection Table Schema	96
127	Colorectal Cancer Pathology Form	96
128	Colorectal Cancer Pathology Table Schema	96
129	Colorectal Cancer Follow-Up Form	97
130	Colorectal Cancer Follow-Up Table Schema	98

131	Breast Cancer Screening Form	99
132	Breast Cancer Screening Table Schema	100
133	Breast Cancer Staging Form	101
134	Breast Cancer Staging Table Schema	102
135	Breast Cancer Resection Form	103
136	Breast Cancer Resection Table Schema	104
137	Breast Cancer Chemotherapy Form	104
138	Breast Cancer Chemotherapy Table Schema	105
139	Breast Cancer Radiotherapy Form	106
140	Breast Cancer Radiotherapy Table Schema	107
141	Breast Cancer Metastatic Treatment Form	108
142	Breast Cancer Metastatic Treatment Table Schema	108
143	Breast Cancer Follow-Up Form	109
144	Breast Cancer Follow-Up Table Schema	110
145	Breast Cancer Pathology Form	110
146	Breast Cancer Pathology Table Schema	111
147	Tumor Size - Accuracy Results (Percentage)	130
148	Tumor Size - Accuracy Results (Percentage) - AdaBoostM1	131
149	Tumor Size - Accuracy Results (Percentage) - Bagging	131
150	Tumor Size - Results Graph	132
151	Tumor Size - Results Graph - AdaBoostM1	133
152	Tumor Size - Results Graph - Bagging	134
153	Tumor Size - T-Test vs. Logistic Regression	134
154	Tumor Size - T-Test vs. Logistic Regression - AdaBoostM1	135
155	Tumor Size - T-Test vs. Logistic Regression - Bagging	135
156	T-Stage - Accuracy Results (Percentage)	136
157	T-Stage - Accuracy Results (Percentage) - AdaBoostM1	136

158	T-Stage - Accuracy Results (Percentage) - Bagging	137
159	T-Stage - Results Graph	137
160	T-Stage - Results Graph - AdaBoostM1	138
161	T-Stage - Results Graph - Bagging	139
162	T-Stage - T-Test vs. Logistic Regression	139
163	T-Stage - T-Test vs. Logistic Regression - AdaBoostM1	140
164	T-Stage - T-Test vs. Logistic Regression - Bagging	140
165	N-Stage - Accuracy Results (Percentage)	141
166	N-Stage - Accuracy Results (Percentage) - AdaBoostM1	141
167	N-Stage - Accuracy Results (Percentage) - Bagging	142
168	N-Stage - Results Graph	142
169	N-Stage - Results Graph - AdaBoostM1	143
170	N-Stage - Results Graph - Bagging	144
171	N-Stage - T-Test vs. Logistic Regression	144
172	N-Stage - T-Test vs. Logistic Regression - AdaBoostM1	145
173	N-Stage - T-Test vs. Logistic Regression - Bagging	145
174	Vascular Involvement - Accuracy Results (Percentage)	146
175	Vascular Involvement - Accuracy Results (Percentage) - AdaBoost M1 $\ .$	146
176	Vascular Involvement - Accuracy Results (Percentage) - Bagging	146
177	Vascular Involvement - Results Graph	147
178	Vascular Involvement - Results Graph - AdaBoostM1	148
179	Vascular Involvement - Results Graph - Bagging	149
180	Vascular Involvement - T-Test vs. Logistic Regression	149
181	Vascular Involvement - T-Test vs. Logistic Regression - AdaBoost M1 $\ .$	150
182	Vascular Involvement - T-Test vs. Logistic Regression - Bagging	150
183	Histology - Accuracy Results (Percentage)	151
184	Histology - Accuracy Results (Percentage) - AdaBoostM1	151

185	Histology - Accuracy Results (Percentage) - Bagging	152
186	Histology - Results Graph	152
187	Histology - Results Graph - AdaBoostM1	153
188	Histology - Results Graph - Bagging	154
189	Histology - T-Test vs. Logistic Regression	154
190	Histology - T-Test vs. Logistic Regression - AdaBoostM1	155
191	Histology - T-Test vs. Logistic Regression - Bagging	155
192	Malignancy - Accuracy Results (Percentage)	156
193	Malignancy - Accuracy Results (Percentage) - AdaBoostM1	156
194	Malignancy - Accuracy Results (Percentage) - Bagging	156
195	Malignancy - Results Graph	157
196	Malignancy - Results Graph - AdaBoostM1	158
197	Malignancy - Results Graph - Bagging	159
198	Malignancy - T-Test vs. Logistic Regression	159
199	Malignancy - T-Test vs. Logistic Regression - AdaBoostM1	160
200	Malignancy - T-Test vs. Logistic Regression - Bagging	160
201	ECOG 6-Month - Accuracy Results (Percentage)	161
202	ECOG 6-Month - Accuracy Results (Percentage) - AdaBoostM1	161
203	ECOG 6-Month - Accuracy Results (Percentage) - Bagging	161
204	ECOG 6-Month - Results Graph	162
205	ECOG 6-Month - Results Graph - AdaBoostM1	163
206	ECOG 6-Month - Results Graph - Bagging	164
207	ECOG 6-Month - T-Test vs. Logistic Regression	164
208	ECOG 6-Month - T-Test vs. Logistic Regression - AdaBoostM1	165
209	ECOG 6-Month - T-Test vs. Logistic Regression - Bagging	165
210	ECOG 9-Month - Accuracy Results (Percentage)	166
211	ECOG 9-Month - Accuracy Results (Percentage) - AdaBoostM1	166

212	ECOG 9-Month - Accuracy Results (Percentage) - Bagging	166
213	ECOG 9-Month - Results Graph	167
214	ECOG 9-Month - Results Graph - AdaBoostM1	168
215	ECOG 9-Month - Results Graph - Bagging	169
216	ECOG 9-Month - T-Test vs. Logistic Regression	169
217	ECOG 9-Month - T-Test vs. Logistic Regression - AdaBoost M1 $\ .$	170
218	ECOG 9-Month - T-Test vs. Logistic Regression - Bagging	170
219	ECOG 12-Month - Accuracy Results (Percentage)	171
220	ECOG 12-Month - Accuracy Results (Percentage) - AdaBoost M1 $\ . \ . \ .$.	171
221	ECOG 12-Month - Accuracy Results (Percentage) - Bagging	172
222	ECOG 12-Month - Results Graph	172
223	ECOG 12-Month - Results Graph - AdaBoostM1	173
224	ECOG 12-Month - Results Graph - Bagging	174
225	ECOG 12-Month - T-Test vs. Logistic Regression	174
226	ECOG 12-Month - T-Test vs. Logistic Regression - AdaBoost M1 $\ \ \ldots \ \ldots$.	175
227	ECOG 12-Month - T-Test vs. Logistic Regression - Bagging	175
228	Survival - Accuracy Results (Percentage)	176
229	Survival - Accuracy Results (Percentage) - AdaBoostM1	176
230	Survival - Accuracy Results (Percentage) - Bagging	176
231	Survival - Results Graph	177
232	Survival - Results Graph - AdaBoostM1	178
233	Survival - Results Graph - Bagging	179
234	Survival - T-Test vs. Logistic Regression	179
235	Survival - T-Test vs. Logistic Regression - AdaBoostM1	180
236	Survival - T-Test vs. Logistic Regression - Bagging	180
237	Tumor Size - R-Squared Results	181
238	Tumor Size - R-Squared Results - AdaBoostM1	181

239	Tumor Size - Regression Results Graph	182
240	Tumor Size - Regression Results Graph - Bagging and Stacking	183
241	Tumor Size - T-Test vs. Linear Regression	183
242	Tumor Size - T-Test vs. Linear Regression - Meta-learners	184
243	ECOG 6-Month - R-Squared Results	185
244	ECOG 6-Month - R-Squared Results - AdaBoostM1	185
245	ECOG 6-Month - Regression Results Graph	186
246	ECOG 6-Month - Regression Results Graph - Bagging and Stacking	187
247	ECOG 6-Month - T-Test vs. Linear Regression	187
248	ECOG 6-Month - T-Test vs. Linear Regression - Meta-learners	188
249	ECOG 9-Month - R-Squared Results	189
250	ECOG 9-Month - R-Squared Results - AdaBoostM1	189
251	ECOG 9-Month - Regression Results Graph	190
252	ECOG 9-Month - Regression Results Graph - Bagging and Stacking	191
253	ECOG 9-Month - T-Test vs. Linear Regression	191
254	ECOG 9-Month - T-Test vs. Linear Regression - Meta-learners	192
255	ECOG 12-Month - R-Squared Results	193
256	ECOG 12-Month - R-Squared Results - AdaBoostM1	193
257	ECOG 12-Month - Regression Results Graph	194
258	ECOG 12-Month - Regression Results Graph - Bagging and Stacking $\ . \ . \ .$	195
259	ECOG 12-Month - T-Test vs. Linear Regression	195
260	ECOG 12-Month - T-Test vs. Linear Regression - Meta-learners	196
261	Survival - R-Squared Results	197
262	Survival - R-Squared Results - AdaBoostM1	197
263	Survival - Regression Results Graph	198
264	Survival - Regression Results Graph - Bagging and Stacking	199
265	Survival - T-Test vs. Linear Regression	199

266	Survival - T-Test vs. Linear Regression - Meta-learners	200
267	Classification - Histology - Data Set C - Bayesian Net 2-Parent	204
268	Classification - Histology - Data Set C - Bayesian Net 2-Parent (continued) .	205
269	Classification - Histology - Data Set C - Confusion Matrix	206
270	Classification - Histology - Data Set C - Joint Probability Distribution Examples	207
271	Classification - Survival - Data Set C - Bayesian Net 2-Parent	216
272	Classification - Survival - Data Set C - Bayesian Net 2-Parent (continued)	217
273	Classification - Survival - Data Set C - Confusion Matrix	218
274	Regression - ECOG 6-Month - Data Set F - Linear Regression w/ Bagging $% \mathcal{A} = \mathcal{A} = \mathcal{A}$.	225
275	Regression - ECOG 9-Month - Data Set F - Multi-layer Perceptron w/ 2	
	Hidden Layers	235

List of Tables

1	QoL/Karnofsky Scores	4
2	ECOG Scores	4
3	Pancreatic Cancer T-Staging	1
4	Pancreatic Cancer N-Staging	1
5	Pancreatic Cancer M-Staging 1	1
6	Tumor Size Distribution	5
7	T-Stage Distribution	5
8	N-Stage Distribution	5
9	Vasculature Involvement Distribution	5
10	Histology Distribution	6
11	Malignancy Distribution	6
12	ECOG 6-Month Distribution	6
13	ECOG 9-Month Distribution	6

14	ECOG 12-Month Distribution	117
15	Survival Distribution	117
16	Classification Experiments	125
17	Regression Experiments	125
18	Classification Algorithms	126
19	Classification Target Values	127
20	Classification Data Sets	127
21	Regression Algorithms	128
22	Regression Experiments	128
23	Regression Data Sets	129
24	Histology Feature-Selected Attribute Subset	203
25	Survival Feature-Selected Attribute Subset	215
26	ECOG 6-Month Feature-Selected Attribute Subset	224
27	ECOG 9-Month Feature-Selected Attribute Subset	234

1 Introduction

The pursuit of cancer research has become one of the most important scientific endeavors of the 21st century. The Cancer Genome Project defines cancer research as *"the intense scientific effort to understand the development of cancer and identify potential therapies"* [Ins]. In 2004, the American Cancer Society announced that cancer had officially replaced heart disease as the highest disease-related cause of death for Americans under the age of 85. Over 1.3 million new cancer cases occurred in the United States in 2005, and it is estimated that one out of every three Americans will be affected by some form of cancer in their lifetime [Soc].

Most major life science fields are already involved extensively in the field of cancer research. Biology and medical science have been an integral part of cancer study since the time of the Ancient Greeks. However, as technologies and therapies evolve in the modern era, there is an increasing demand for specialized advances from the field of computer science. Just a few of computer science's contributions to cancer research include diagnostic tools, predictive modeling, imaging and data analysis, bioinformatics, medical training applications, and collaborative research databases. Discoveries from computer science are already implemented in a wide variety of cancer therapies, including surgery, radiotherapy, chemotherapy, diagnostic imaging, immunotherapy, and genetic therapy.

Study of clinical performance is one of cancer research's most important research subjects, as it directly concerns the patient's wellbeing. Clinical performance refers to a patient's response to applied medical therapy. Response factors may include changes in health, progression of illness, disease pathology, and systemic behaviors of the body. More refined analysis of clinical performance is always needed, given the frequent complexity and difficulty of cancer treatment. These analyses may include building predictive models for clinical performance generated using the data mining and machine learning techniques from the field of computer science.

Our goal in this research is twofold: to develop clinical performance databases of cancer patients, and to conduct data mining and machine learning studies on the collected patient records. We present a novel database designed by UMass Memorial Medical School oncologists for representing highly-detailed clinical performance of breast and gastrointestinal cancer patients. Machine learning techniques will be applied to the patient contents of this database to generate a variety of predictive models. The tools and techniques of data mining and machine learning are ideal for this type of analysis. We present and evaluate models based on pancreatic cancer patient data for predicting disease characteristics and prognosis of survival and wellbeing.

This research is a joint effort between the WPI Computer Science Department and UMass Memorial Medical School. The clinical database is composed of data from patients seen at the UMass Memorial Department of Surgical Oncology. This project is advised by Prof. Carolina Ruiz, whose research focus is machine learning and data mining. Prof. George Heineman of WPI and Prof. Sergio Alvarez of Boston College provided additional computer science advising. Medical advising is provided by the Surgical Oncology staff at UMass Memorial, particularly Dr. Giles Whalen and Mary Sullivan NP for the gastrointestinal module, and Dr. Robert Quinlan for the breast module. A grant provided by UMass Memorial in August 2005 funded this research.

2 Medical Background

Cancer refers to diseases resulting from uncontrolled cell growth in regions known as *neo*plasms or tumors. A tumor may refer to any distinct mass in a tissue or organ, and its growth may either be benign or malignant. Malignant tumors are characterized by their ability to spread to surrounding local tissue (*invasion*) or distant sites in the body (*metastasis*). The malignant tumors discussed in this research are a form of cancer known as *carcinoma*, or cancers arising from *epithelial* cells. Tumor growth may be caused by damage or mutations to cell DNA from different factors, including hereditary conditions, environmental exposure, and infectious disease. Chemical or physical agents which trigger cancer-causing DNA mutations are referred to as *carcinogens*. Symptoms of cancer depend on the site of the body affected, the nature of the tumor, and metastatic spread of the disease.

Oncology is the branch of medicine which deals with the diagnosis and treatment of malignant tumors. Various methods exist to treat cancer. Resection is the surgical excision of tumor growth from bodily tissue. Chemotherapy is the systemic or localized application of antineoplastic drugs to destroy or retard the development of tumor growth. Radiotherapy refers to treatments which use irradiation to destroy cancerous cells. Palliation collectively refers to the methods intended to relieve cancer symptoms rather than effect cure. Palliative measures may include stenting, anastomosis, feeding tubes, nerve blocks, and various forms of surgery, chemotherapy, and radiotherapy, as well as other medications for symptom management. The intention of a resection may be either curative or palliative. Tumor immunotherapy is a biological protocol which uses methods such as vaccination to trigger an immune system response which destroys cancerous cells. Gene counseling is a series of DNA tests which establish susceptibility of a patient or their family to certain forms of cancer.

An important aspect of patient clinical performance research is quantification of a patient's wellbeing. Measurements of wellbeing are important in evaluating treatment response and qualifications for different forms of care. Throughout the course of their treatment, patient overall health and performance status may be rated by quality-of-life (QoL) scores

Score	Status
100%	Normal, No Complaints, No Signs of Disease
90%	Capable of Normal Activity, Few Symptoms or Signs of Disease
80%	Normal Activity with Some Difficulty, Some Symptoms or Signs
70%	Caring for Self, Not Capable of Normal Activity or Work
60%	Requiring Some Help, Can Take Care of Most Personal Requirements
50%	Requires Help Often, Requires Frequent Medical Care
40%	Disabled, Requires Special Care and Help
30%	Severely Disabled, Hospital Admission Indicated but No Risk of Death
20%	Very Ill, Urgently Requiring Admission, Requires Treatment
10%	Moribund, Rapidly Progressive Fatal Disease Processes
0%	Death

Table 1: QoL/Karnofsky Scores

Score	Status	
0	Asymptomatic	
1	Symptomatic but Completely Ambulant	
2	Symptomatic, $<50\%$ in Bed During the Day	
3	Symptomatic, $>50\%$ in Bed, but Not Bedbound	
4	Bedbound	
5	Death	

Table 2: ECOG Scores

(also known as Karnofsky scores), which ranges 0-100%, or Eastern Cooperative Oncology Group (ECOG) scores, which ranges 0-5. Tables 1 and 2 detail the criteria for these scores [KB49, OC82]. For the purpose of this thesis, patient wellbeing will be measured using the ECOG system.

Different factors may be used to describe the nature of tumors. *Histology* refers to the microscopic structure of tumor tissue. The behavior and severity of a cancer may vary depending on its histologic composition. *Adenocarcinoma* is carcinoma which develops within glandular epithelium which typically behaves in a very malignant fashion. *Neuroendocrine* tumors grow in nervous or endocrine tissue. For some cancers, including malignancies of the pancreas, these neuroendocrine tumors tend to behave in a more indolent fashion than adenocarcinomas. *Cysts* refer to closed cavities of glandular epithelium where retained se-

cretions are accumulated, and may behave in a benign or malignant fashion. Two common histologic forms of breast cancer are *lobular* and *ductal* types. The study of cells at a microscopic level is referred to as *cytology*. At the microscopic level, the symptoms of cancer are often influenced by the growth and penetration of tumors into bodily structures. *Lymph nodes* are small bodies along lymphatic vessels which filter bacteria and foreign bodies. The presence of tumorous tissue within regional lymph nodes is an important prognostic factor for many types of cancer. The penetration of tumors into *vasculature*, or blood vessels, can be an important factor in determining the spread and resectability of the disease.

The American Joint Committee on Cancer (AJCC) maintains a staging system to provide a unified methodology for describing cancer. Malignant tumors are classified by TNM staging, which refers to Tumor, Node, and Metastasis. Each parameter is paired with a number from a discrete range to indicate disease stage. The meaning of these parameters differs by cancer etiology. T refers to primary tumor size and ranges from 0 to 4 or 'is' for *in situ* growth. N refers to regional lymph node involvement and ranges from 0 to 3. M refers to metastatis to distant organs and is denoted 0 if absent and 1 if present. Other parameters may be used to describe cancer. R is used to denote tumor growth on margins of surgically excised tissue: 0 for clean margins, 1 for microscopic tumor growth, and 2 for gross tumor growth. L and V (0-1) denote the absence or presence of tumor invasion into lymphatic vessels and veins. G (1-4) stands for the grade or differentiation between tumor cells and surrounding normal cells. The criteria for staging depends on the tumor location and histology. Most tumor forms use TNM staging, but not all use the full range. In all staging systems, a parameter paired with X stands for an unknown or unevaluated quantity [oC04].

A variety of tools are used to diagnose cancer. *Serum studies* refer to blood tests, which may include nutritional levels, liver functions, and molecular *tumor markers*. *Biopsy* refers to a small sample of tumor tissue taken to evaluate its histologic composition and malignancy. Biopsies may be taken in a variety of ways, including fine-needle aspiration (FNA), corecutting needle, incisional biopsy, and excisional biopsy. Cancer is frequently diagnosed using *imaging studies*. Quantifying the accuracy and reliability of imaging studies is a crucial research topic. X-rays are the process of visualizing an internal body image by catching highenergy photons on photographic film. A *computed axial tomography* (CT or CAT) creates a three-dimensional internal view of a patient using a series of sectional x-rays across a common axis. Ultrasound uses ultrasonic waves to create a sonographic visualization a body's internal structure. Endoscopic ultrasound (EUS) is an ultrasound study generated by a thin, flexible ultrasound probe passed through the gastrointestinal tract. Magnetic resonance imaging (MRI) uses the magnetic resonance of photons to create a high-contrast density image. Biopsies are often taken using guidance by imaging studies. Different diagnoses are used depending on the type and location of cancer [VD93].

2.1 Pancreatic Cancer Background

Pancreatic cancer remains a challenging disease for physicians, oncologists, and surgeons, and is the machine learning analytic focus of this thesis. Here, pancreatic cancer is a general term for cancer of the pancreas and periampullary region. The *pancreas* is a long gland which sits behind the stomach and secretes digestive juices into the small intestine and bloodstream. The *periampullary region* refers to the area containing the duodenum, distal common bile duct, and ampulla of Vater. The *duodenum* refers to the upper part of the small intestine, which starts from the lower end of the stomach and extends to the *jejunum* (middle small intestine). The *distal common bile duct* is the portion of the excretory passage close to the duodenum which carries bile from the liver. The ampulla of Vater is a dilation in the duodenal wall through which the common bile duct and pancreatic duct empty into the small intestine. Please refer to Figures 1 and 2 [Gra95, Cen].

Tumors of the pancreatic and periampullary region are known for a high degree of mortality and morbidity. This disease stands as the fourth largest cancer killer in the country, even though it only accounts for 2% of total cancer diagnoses. Approximately 25,000 new patients are diagnosed with this disease in the United States each year; median survival from time of diagnosis is six months, with five-year survival rates at 3% [Bre04]. The severity and treatment of these cancers depend largely on their locations and histologic types. The most frequently occurring types are adenocarcinomas, which are the most aggressive and have the highest associated mortality rates. A less common and more indolent form of the disease are neuroendocrine or *islet cell* tumors. *Intraductal papillary mucinous neoplasms* (IPMNs or IPMT's) are cystic pancreatic tumors which can progress to cancers.



Figure 1: Gray's Anatomy - Pancreas and Periampullary Region [Gra95]



Figure 2: Digestive System with Common Bile Duct Illustrated [Cen]

Pancreatic cancer typically presents itself through non-specific symptoms, abdominal pain and painless jaundice being the most frequent. Risk factors include age, smoking, obesity, diabetes, diets high in meat, chronic pancreatitis, and genetic family history. Diagnosis is typically performed using chest x-rays, serum studies, abdominal CT scans, and endoscopic ultrasound. Imaging studies be used to determine tumor size, regional lymph note involvement, and distant metastatic spread. Biopsies taken by *fine needle aspiration* (FNA) during endoscopic ultrasound can be used to predict tumor histology and malignancy. Nuclear tumor markers such as CEA and CA19-9, as well as nutritional and liver function serum levels, can confirm the systemic presence of pancreatic cancer or evaluate its effects. In preliminary evaluation, approximately 15% of patients are deemed as potentially resectable, 40% as locally advanced/unresectable, and 45% as metastatic or equivocal.

TNM staging for pancreatic cancer determines the treatment course and prognosis of disease. The T-stage in pancreatic cancer refers to the tumor's size and penetration into surrounding gastrointestinal anatomy. A simplified version of the AJCC staging criteria [oC04] is presented in Table 3. Regional lymph node involvement as denoted by N-stage and presence of metastatis as denoted by M-stage is presented in Tables 4 and 5. Tumor spread in pancreatic cancer may involve vascular structures, which impacts disease spread and difficulty of resection. Vascular structures which may be invaded include the *celiac axis*, *hepatic artery*, *superior mesenteric artery*, *superior mesenteric vein*, *inferior vena cava*, *portal vein*, and *splenic vein*. If a tumor penetrates a venous structure, then sections of the vein may be resected. However, arterial penetrations cannot be resected given current medical technology, although studies are being done. The microscopic penetration of tumor into a vascular structure is denoted by V-staging as described above.

The most common surgical procedure to treat pancreatic cancer is a *Whipple procedure*, or *pancreaticoduodenectomy*. The procedure involves removal of the distal half of stomach, gall bladder, distal common bile duct, head of the pancreas, duodenum, proximal jejunum, and regional lymph nodes. The remaining anatomy is anastomosed together to reconstruct

T-Stage	Criteria			
TX	Primary tumor cannot be assessed			
Т0	No evidence of primary tumor			
Tis	Carcinoma in situ			
T1	Tumor limited to pancreas and measures 2 cm or			
	less in greatest dimension, without blood vessel involvement			
Τ2	Tumor greater than 2 cm in greatest dimension, still			
	limited to the pancreas, without involve any blood vessels			
Т3	Any tumor that extends beyond the pancreas, does not			
	involve the celiac axis or superior mesenteric artery.			
Τ4	Any tumor that invades the superior mesenteric artery			
	or the celiac axis (unresectable cancer)			

Table 3: Pancreatic Cancer T-Staging

N-Stage	Criteria
NX	Regional lymph node involvement cannot be assessed
NO	No evidence of regional lymph node involvement
N1	Presence of regional lymph node involvement

Table 4: Pancreatic Cancer N-Staging

M-Stage	Criteria	
MX	Distant metastasis cannot be assessed	
M0	No evidence of distant metastasis	
M1	Presence of distant metastasis	

Table 5: Pancreatic Cancer M-Staging

a working digestive tract. The pre and post-surgical anatomy of a Whipple procedure are shown in Figures 3 and 4 [Cli]. The surgical mortality rate of a Whipple procedure is approximately 5%, 3% in high-volume centers. Resective surgery is usually performed in most circumstances where possible, as it represents the highest likelihood of complete cure. Reasons not to resect include local tumor spread, involvement of vasculature, distant metastatis, and patient unwillingness or inability to endure surgery.

Chemotherapy and radiotherapy are frequently applied as pancreatic cancer treatments. The most common regimens of chemotherapy applied at UMass Memorial are 5-Flurouracil and Gemcitabine. Cancer therapies may be either *adjuvant* (applied post-surgery) or *neoad-juvant* (applied pre-surgery, frequently in an effort to reduce tumor size). Palliative measures intended to alleviate but not cure disease include feeding tubes, stenting, gastric bypasses, nerve blocks, and palliative chemo or radiotherapy. After initial treatment, patients are followed at three-month intervals for the first two years, and six-month intervals for two to five years, and yearly intervals afterwards. Factors monitored during follow-up include disease status, recurrent symptoms, weight, serum markers, and general wellbeing scores.



Figure 3: Whipple Procedure - Pre-Surgical Anatomy [Cli]



Figure 4: Whipple Procedure - Post-Surgical Anatomy [Cli]

3 Clinical Database Construction

The clinical database is where our patient information is collected. Our database was developed using Microsoft Access 2003 with Visual Basic scripting and SQL Server for data storage. It is hoped that these additional cancer modules will be used in future analytic work. Prof. George Heineman of WPI and [Szo82] provided many useful suggestions in representing the patient treatment narrative within a software application.

Specific details pertaining to the patient medical factors are too complex to be discussed here; for those interested, [VD93] provides an accessible discussion of clinical oncology for both medical and non-medical audiences alike.

3.1 Gastrointestinal Cancer Database

For this project, database modules were developed for six major forms of gastrointestinal cancer (pancreatic, biliary, esophageal, gastric, colorectal, and hepatocellular). Specific design of the gastrointestinal cancer modules were based on Dr. Whalen algorithms for patient treatment. Portions of the table schema and interface were based on earlier work by Tiffany Wei of UMass Memorial.

In this database, the major elements of patient treatment were decomposed into eight categories:

- Presentation
- Medical History
- Diagnostic Tests
- Preliminary Outlook
- Treatment
- Surgical Resection Details/Reasons for Not Pursuing Resection

- Pathology Reports
- Follow-Up

Each of these categories is represented by a table schema within the database. They are related to a core patient record by a zero-to-many cardinality; this allows for a flexible, efficient representation of what can often be a very complex clinical narrative.

3.1.1 Pancreatic Cancer

		Window Help Adobe		- 8 Q.
MS Sans Serif	→ 8 → B		<u>∆</u> • <u>∧</u> • <u>∕</u> • □	·
11		a a 280 Taes sea sea		•
resumptive Diagn	osis at Onset of	Care		
	*			
resentation				
Date of Evaluation ECOG	Performance Status	Height (in.) Weig	it (lbs.)	
© 0 0				
<u>Symptoms</u>				
<u>Symptoms</u> I≣ Weight Loss	🕅 Biliary Colic	🕅 Pruritis		
🔟 Weight Loss	🕅 Biliary Colic	👿 Pruritis 🗑 Abdominal Pain	Imigestion I∭ Dysphagia	
How Much (pounds):	🕅 🕅 Nausea	I≣ Abdominal Pain I≣ Back Pain	🕅 Dysphagia	
 Weight Loss How Much (pounds): Jaundice 	│ IIII Nausea IIII Vomiting	I≣ Abdominal Pain I≣ Back Pain	🕅 Dysphagia	

Figure 5: Pancreatic Cancer Presentation Form

Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)	
DemEvalDate	Date/Time	Demographics - Date Evaluated by Surgical Oncology	
DemECOG	Number	Demographics - ECOG Score (0-4)	
DemHeight	Number	Demographics - Height in Inches of Patient	
DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission	
SxWtloss	Yes/No	Initial Symptoms - Weight Loss	
SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds	
SxJaun	Yes/No	Initial Symptoms - Juandice	
SxChole	Yes/No	Initial Symptoms - Cholecystitis	
SxChola	Yes/No	Initial Symptoms - Cholangitis	
SxBC	Yes/No	Initial Symptoms - Biliary Colic	
SxNau	Yes/No	Initial Symptoms - Nausea	
SxVom	Yes/No	Initial Symptoms - Vomiting	
SXCCS	Yes/No	Initial Symptoms - Clay Colored Stool	
SxFati	Yes/No	Initial Symptoms - Fatigue	
SxPru	Yes/No	Initial Symptoms - Pruritis	
SxInd	Yes/No	Initial Symptoms - Indigestion	
SxAbd	Yes/No	Initial Symptoms - Abdominal Pain	
SxBack	Yes/No	Initial Symptoms - Back Pain	
SxDyspha	Yes/No	Initial Symptoms - Dysphagia	
SxSatiety	Yes/No	Initial Symptoms - Early Satiety	
SxOT	Yes/No	Initial Symptoms - Other	
SxOTSpe	Text	Initial Symptoms - Other - Specify	

Figure 6: Pancreatic Cancer Presentation Schema

Microsoft Access - [Pan_2	_History : Form]		-OX
🗐 File Edit View Insert	Format <u>R</u> ecords <u>T</u> ools	Window Help Adobe PDF	_ 8 ×
⊻ • = € 5 0 . ♥			0.
Medical History			
	Malnutrition Liver Failure/Cirrhosis Diabetes Less than Six Months Greater than Six Months Oral Agents Diet Control ar Drug Use	Cancer History Patient Prior Dx: Im Chemo Im Chemo Radiation Surgery Father Dx: Im Chemo Mother Dx: Im Chemo Related Dx: Other Related Dx: Im Chemo Im Chemo	
Form View			

Figure 7: Pancreatic Cancer Medical History Form

Field Name Data Type		Description			
1D	AutoNumber	ID			
MR	Text	Meditech Medical Record Number for Patient			
CxHF	Yes/No	Comorbidities - Heart Failure			
CxIHD	Yes/No	Comorbidities - Ischemic Heart Disease			
CxResp	Yes/No	Comorbidities - Respiratory			
CxDiab	Yes/No	Comorbidities - Diabetes			
CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral			
CxDiabDiet	Yes/No	Comorbidities - Diabetes - Insulin - Diet Control			
CxDiabOnset	Number	Comorbidities - Diabetes - Onset (1=Less than six months, 2 =Greater than six months)			
CxRF	Yes/No	Comorbidities - Renal Failure			
CxHyper	Yes/No	Comorbidities - Hypertension			
CxBleed	Yes/No	Comorbidities - Bleeding Disorder			
CxLiver	Yes/No	Comorbidities - Liver Failure			
CxMal	Yes/No	Comorbidities - Malnutrition			
CxPriorCancer	Number	Comorbidities - Prior Cancer Dx			
CxPriorCancerChemo Yes/No		Comorbidities - Prior Cancer Dx - Chemo			
CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation			
CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery			
SHCigarette	Yes/No	Social History - Cigarettes (significant use)			
SHAlcohol	Yes/No	Social History - Alcohol (significant use)			
SHDrugUse	Yes/No	Social History - Drug Use			
SHExposure	Yes/No	Social History - Environmental Exposure			
SHOther	Yes/No	Social History - Other			
SHOtherS	Text	Social History - Other - Specify			
FamilyFatherDx	Number	Family History - Father Dx			
FamilyMotherDx	Number	Family History - Mother Dx			
FamilyOther1	Text	Family History - Other1			
FamilyOther1Dx	Number	Family History - Other1 Dx			
FamilyOther2	Text	Family History - Other2			
FamilyOther2Dx	Number	Family History - Other2 Dx			

Figure 8: Pancreatic Cancer Medical History Table Schema
EB File	e <u>E</u> dit	View Insert	Format	Reco) 문화	1000		파란성	Ado <u>b</u> e F	PDF		- 8	×
₩ •		a 🕽 💞	X 电	B	0		V.	10	7 🐴	>* 1%	₫ /a •	2.	
	,	MS Sans Ser	f	- 8	3 +	BI	U		E 30	2 - 1	 1 -	•	>
Serur	n Stuc	lies									-		
CEA:		Albumi	n [AL	.K			ALT				
CA19-9:		Total B	ilirubin [AS	6T:			Amyla	se:			

Figure 9: Pancreatic Cancer Serums Studies Form

Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
LabCEA	Number	Laboratory - CEA	
LabCA19-9	Number	Laboratory - CA19-9	
LabAlb	Number	Laboratory - Albumin	
LabBili	Number	Laboratory - Bilirubin	
LabAlka	Number	Laboratory - Alkaline phosphotase	
LabALT	Number	Laboratory - ALT	
LabAST	Number	Laboratory - AST	
LabAmylase	Number	Laboratory - Amylase	
1 N			

Figure 10: Pancreatic Cancer Serums Studies Table Schema

⊻ - 12 	3 0 % 2↓ X↓ Vy Ta ∇ M >+ M B □ a • 0 .
iagnostic Imaging Procedur	es
CT with Panreatic Protocol/CT/	<u>x</u>
Date of Procedure:	
🕅 Celiac Artery Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown
🕅 Superior Mesenteric Artery Involvement	© Open © Abutted © Encased © Occluded © Unknown
👿 Hepatic Artery Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown
🕅 Inferior Vena Cava Involvement	Copen C Abutted C Encased C Occluded C Unknown
Superior Mesenteric Vein Involvement	Copen C Abutted C Encased C Occluded C Unknown
Portal Vein Involvement	Copen C Abutted C Encased C Occluded C Unknown
Nodes: 🕅 Celiac Nodal Disease	The Other Nodal Disease The No Nodal Assessment or Mention
Tumor Size (cm): by	
Chest X-Ray (CXR)	
Percutaneous Transhepatic Cl	iolangiography (PTC)
Date of Procedure:	
Stenting Type: @ Internal	le External

Figure 11: Pancreatic Cancer Diagnostic Imaging Form

	Field Name	Data Type	Description		
		AutoNumber	ID		
1	1R	Text	Meditech Medical Record Number for Patient		
	XRDx	Yes/No	CXR - Diagnosis		
	TDx	Yes/No	CT - Diagnosis		
	TEvalDate	Date/Time	CT - Date Evaluated		
	TVascOmit	Yes/No	CT - Vascular Omission		
	TCeliac	Yes/No	CT - Celiac Involvement		
	TCeliacClass	Number	CT - Celiac Involvement Class		
	TSMA	Yes/No	CT - SMA Involvement		
	TSMAClass	Number	CT - SMA Involvement Class		
CTHepatic Yes/No CT - Hepatic Involvement			CT - Hepatic Involvement		
	CTHepaticClass Number CT - Hepatic Involvement				
	TInferior	Yes/No	(No CT - Inferior Vena Cava Involvement		
	TInferiorClass	Number	CT - Inferior Vena Cava Involvement Class		
	TSMV	Yes/No	CT - SMV Involvement		
	TSMVClass	Number	CT - SMV Involvement Class		
	TPortal	Yes/No	CT - Portal Vein Involvement		
	TPortalClass	Number	CT - Portal Vein Involvement Class		
	TCeliacNode	Yes/No	CT - Celiac Nodal Disease		
	TOtherNode	Yes/No	CT - Other Nodal Disease		
	TNodeOmit	Yes/No	CT - Node Omission		
	TTumorSizeX	Number	CT - Tumor Size (cm) - Width		
	TTumorSizeY	Number	CT - Tumor Size (cm) - Height		
F	TCDx	Yes/No	PTC - Diagnosis		
F	'TCEvalDate	Date/Time	PTC - Date Evaluated		
F	TCStent	Yes/No	PTC - Stent		
F	TCStentType	Number	PTC - Stent Type		

Figure 12: Pancreatic Cancer Diagnostic Imaging Table Schema

Microsoft Access - [Pan_3c_Endoscopy	/:Form]	- 🗆 ×
🖪 File Edit View Insert Format !	Records Tools Window Help Adobe PDF Type a question for help	8 ×
1 K • H • 6 6 V X • 1		
	· · · · · · · · · · · · · · · · · · ·	÷.
Endoscopy Procedures		
Endoscopic Ultrasound (EUS)		
Date of Procedure:		
🕅 Celiac Artery Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown	
🕅 Superior Mesenteric Artery Involvement	🕫 Open 🕼 Abutted 🕼 Encased 🕼 Occluded 🕼 Unknown	
🕅 Hepatic Artery Involvement	Ø Open Ø Abutted Ø Encased Ø Occluded Ø Unknown	
🔟 Inferior Vena Cava Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown	
🕅 Superior Mesenteric Vein Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
🕅 Portal Vein Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown	
Nodes: 🖉 Celiac Nodal Disease	🕅 Other Nodal Disease 📓 No Nodal Assessment or Mention	
Tumor Size (cm): by		
EUS Staging: T 💽 N	FNA Cytology	
Endoscopic Retrograde Chola	ngiopancreatogram (ERCP)	
Date of Procedure:		
Stenting Type: Plastic	Metal	
Form View		//

Figure 13: Pancreatic Cancer Endoscopy Studies Form

Field Name	Data Type	Description
)D	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
EUSDx	Yes/No	EUS - Diagnosis
EUSEvalDate	Date/Time	EUS - Date Evaluated
EUSVascOmit	Yes/No	EUS - Omission
EUSCeliac	Yes/No	EUS - Celiac Involvement
EUSCeliacClass	Number	EUS - Celiac Involvement Class
EUSSMA	Yes/No	EUS - SMA Involvement
EUSSMAClass	Number	EUS - SMA Involvement Class
EUSHepatic	Yes/No	EUS - Hepatic Involvement
EUSHepaticClass	Number	EUS - Hepatic Involvement Class
EUSInferior	Yes/No	EUS - Inferior Vena Cava Involvement
EUSInferiorClass	Number	EUS - Inferior Vena Cava Involvement Class
EUSSMV	Yes/No	EUS - SMV Involvement
EUSSMVClass	Number	EUS - SMV Involvement Class
EUSPortal	Yes/No	EUS - Portal Vein Involvement
EUSPortalClass	Number	EUS - Portal Vein Involvement Class
EUSCeliacNode	Yes/No	EUS - Celiac Node Disease
EUSOtherNode	Yes/No	EUS - Other Nodal Disease
EUSNoNode	Yes/No	EUS - No Nodes Mentioned
EUSTumorSizeX	Number	EUS - Tumor Size (cm) - Width
EUSTumorSizeY	Number	EUS - Tumor Size (cm) - Height
EUSStagingT	Number	EUS - Staging - T
EUSStagingN	Number	EUS - Staging - N
EUSCyto	Number	EUS - FNA Cytology
ERCPDX	Yes/No	ERCP - Diagnosis
ERCPEvalDate	Date/Time	ERCP - Date Evaluated
ERCPStent	Yes/No	ERCP - Stent
ERCPStentType	Number	ERCP - Stent Type

Figure 14: Pancreatic Cancer Endoscopy Studies Table Schema

		Format Re	ecords <u>T</u> ools	; <u>W</u> indow	Help	Ado <u>b</u> e F	PDF	Type a qu	estion for he	lp 💌	- 8
🖌 🕶 🔛	a 🛯 🖤	X 电 @	0		1	VM) * H	(😭	🗇 🕼 📲	2.	
	-	¥	*	BZU		F H	۵.	<u>A</u>	2		• .
re-Surgi	cal Outlo	ok									
Potentially	Resectable										
Locally Adv	anced/Unres	ectable									
Metastatic	or Equivocal F	indings									

Figure 15: Pancreatic Cancer Preliminary Outlook Form

6 6	Field Mamo	Data Type	Description
DI	Search	AutoNumber	ID
MR		Text	Meditech Medical Record Number for Patient
PreOu	tlook	Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic or Equivocal Findings)

Figure 16: Pancreatic Cancer Preliminary Outlook Table Schema

😨 Eile Edit View Insert Format Record	ds Iools Window Help Ad ⊃ 🍓 🛃 🕺 🏹 🏹 🎦	lobe PDF Type a question for f	help - 8 >
*	• B I U E E	≡ 2 • 4 • 2 •	• = • •
Treatment Course			
Resection	🖩 Staging Laparos	copy/Laparotomy	
Radiation Adjuvant @ Neoadjuvant @ Both	🖩 Palliative Measu	res	
Chemotherapy	🗰 Bypass 🕅 Gastrostomy Tube	I HAL I PV Shunts	
Adjuvant R Neoadjuvant R Both	🕅 Jejunstomy Tube	🕅 Pall. Stenting	
Avastin 🕅 Leukovorin	🕅 Celiac Block	🕅 Pall. Radiation	
Capecitabine 🕅 Levamasole	🕅 Paracentesis	🕅 Pall. Resection	
🕼 Erbitux 🕼 Mitomycin	🕅 Thoracentesis	Two Other - Specify:	
🐺 Fluorouracil (5-FU) 🕅 Oxaliplatin	🕅 Transfusion		
FUDR Taxol Gencitabine Tother - Specify:	🖩 <u>Experimental Pro</u>	otocol (vaccine, etc.)	
🕅 Irinotecan	Genetic Counsel	ing	
		AND RECK	

Figure 17: Pancreatic Cancer Treatment Form

AutoNumber Text Yes/No Yes/No Yes/No Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	ID Meditech Medical Record Number for Patient Treatment - Resection Treatment - Laparoscopy Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Evolux Treatment - Chemo - Fluorouracil (S-FU) Treatment - Chemo - FUDR Treatment - Chemo - Fuorouracil (S-FU) Treatment - Chemo - Fuorouracil (S-FU)
Yes/No Yes/No Number Yes/No Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Resection Treatment - Laparoscopy Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Evolution Treatment - Chemo - Evolution Treatment - Chemo - Fullor Treatment - Chemo - FUDR Treatment - Chemo - FUDR Treatment - Chemo - Fulcabine
Yes/No Yes/No Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Laparoscopy Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo Treatment - Chemo - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Avastin Treatment - Chemo - Erbitux Treatment - Chemo - Fulorouracii (S-FU) Treatment - Chemo - FUDR Treatment - Chemo - Fulorouracii (S-FU)
Yes/No Number Yes/No Yes/No	Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Avastin Treatment - Chemo - Evolutione Treatment - Chemo - Evolutione Treatment - Chemo - Evolutione Treatment - Chemo - Fluorouracii (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - FUDR Treatment - Chemo - Fuoroania
Number Yes/No Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Radiation - Adjuvancy Treatment - Chemo - Treatment - Chemo - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Fuorouracil (5-FU)
Yes/No Number Yes/No	Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gencitabine Treatment - Chemo - FuDR Treatment - Chemo - Fuccan
Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Exbitux Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracii (S-FU) Treatment - Chemo - FUDR Treatment - Chemo - Fuore
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Avastin Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracii (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Fluorouracii (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gencitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No	Treatment - Chemo - Irinotecan
Yes/No Yes/No	
Yes/No	
	Treatment - Chemo - Leukovorin
	Treatment - Chemo - Levamasole
Testino	Treatment - Chemo - Mitomycin
Yes/No	Treatment - Chemo - Oxaliplatin
Yes/No	Treatment - Chemo - Taxol
Yes/No	Treatment - Chemo - Other
Text	Treatment - Chemo - Other - Specify
Yes/No	Treatment - Palliation
Yes/No	Treatment - Palliation - Pall. Resection
Yes/No	Treatment - Palliation - Bypass
Yes/No	Treatment - Palliation - Celiac Block
	Treatment - Palliation - Paracentesis
	Treatment - Palliation - Thoracentesis
	Treatment - Palliation - Pall, Radiation
	Treatment - Palliation - Transfusion
	Treatment - Palliation - Pall. Stenting
	Treatment - Palliation - PV Shunts
	Treatment - Palliation - HAI
	Treatment - Palliation - Gastrostomy Tube
	Treatment - Palliation - Jejunstomy Tube
	Treatment - Palliation - Other
	Treatment - Palliation - Other - Specify
	Treatment - Experimental protocol (ie. vaccine)
	Treatment - Gene Courseling
	Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No

Figure 18: Pancreatic Cancer Treatment Table Schema

Microsoft Access - [Pan_6a_Res : Form]
🗃 File Edit View Insert Format Records Iools Window Help Adobe PDF Type a question for help 🔹 🗗 🗙
IM - ■ • ● ● · * * ■ ■ · • ● ● ↓ ↓ ▼ • ▼ M · * × @ ● ● · • .
MS Sans Serif • 8 • B I U ≣ ≣ ≣ 2 • ▲ • 2 • □ • .
If Resection is Performed
Surgery Date of Admission: Date of Surgery: Procedure Type
🗑 Venous Resection 🖉 Venous Reconstruction 🕅 Arterial Resection 🖉 Arterial Reconstruction
Other Organs Resected: Estimated Blood Loss (cc):
🕅 Transfusion If Yes, Units: 🚺 Methods: 📓 FFP 📓 Cell Saver
Resection Attempt: @ Successful @ Unsuccessful - Reason: Post-Op Days in ICU: Post-Op Care Path: @ Congruent @ Divergent
ING/Gastrostomy Drainage > 7days I Abdominal Collection
🖩 Pulmonary Complications 🗐 Wound Infection 📓 Leak
🕼 Liver Insufficiency (Total Bilirubin > 5) If Yes, Total Bilirubin:
Date of Discharge: Discharge Status:
Form View

Figure 19: Pancreatic Cancer Resection Form

Field Name	Data Type	Description
[D	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
ResDAdm	Date/Time	Resection - Date of Admission
ResDSurg	Date/Time	Resection - Date of Surgery
ResPxType	Number	Resection - Procedure Type (Whipple, total pancreatectomy, distal pancreatectomy, etc)
ResORTime	Number	Resection - OR Time (hr.)
ResVenRes	Yes/No	Resection - Venous Resection
ResVenRec	Yes/No	Resection - Venous Reconstruction
ResArtRes	Yes/No	Resection - Arterial Resection
ResArtRec	Yes/No Resection - Arterial Reconstruction	
ResOrgans	Text	Resection - Other Organs Resection
ResBloodLoss	Number	Resection - Estimated Blood Loss (cc)
ResTransfusion	Yes/No	Resection - Tranfusion
ResTUnits	Number	Resection - Transfusion Units
ResTFFP	Yes/No	Resection - Transfusion - FFP
ResTCell	Yes/No	Resection - Transfusion - Cell
ResAttempt	Number	Resection - Resection Attempt
ResAttemptUn	Number	Resection - Resection Unsuccessful Reason (Tumor involvement, Operative mishap, etc)
ResPOCourse	Number	Resection - PO - Post-Op Care Path
ResPODays	Number	Resection - PO - Time in ICU (days)
ResPOInfection	Yes/No	Resection - PO - Wound infection
ResPOLeak	Yes/No	Resection - PO - Leak
ResPONG	Yes/No	Resection - PO - NG/gastrotomy drainage
ResPOAbdominal	Yes/No	Resection - PO - Abdominal Collection
ResPOPulmComp	Yes/No	Resection - PO - Pulminary Complications
ResPOLiverInsuf	Yes/No	Resection - PO - Liver Insufficiency
ResPOLiverTB	Number	Resection - PO - Liver Insufficiency - Total Bilirubin
ResPODDischarge	Date/Time	Resection - Date of Discharge
ResPODischStatus	Number	Resection - Discharge Status

Figure 20: Pancreatic Cancer Resection Table Schema

Eile Edit View Insert Format Record	s <u>T</u> ools <u>W</u> indow <u>H</u> elp Ado <u>b</u> e PDF	Type a question for help 🛛 💂 🗗 🕽
Resection is Not Performed	1	
Date of Decision:		
teasons (select all that apply):		
Clinical Decision		
🕅 Patient Couldn't Handle Proposed Treatment		
🕅 Patient Refused Treatment		
Proposed Magnitude of Treatment and Risks Not Worth Likely Benefit		
<u>Vascular Involvement</u>	Additional Disease	
🕅 Celiac Artery Involvement	🕅 Cirrhosis	
🐷 Superior Mesenteric Artery Involvement	Evidence of Metastasis	
🕅 Hepatic Artery Involvement		
👿 Inferior Vena Cava Involvement		
Superior Mesenteric Vein Involvement		
Portal Vein Involvement		
I₩ Portal Vein Involvement		

Figure 21: Pancreatic Cancer No Resection Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
NoResEvalDate	Date/Time	No Resection - Decision Date
NoResNoHandle	Yes/No	No Resection - Couldn't Handle Proposed Treatment
NoResRefused	Yes/No	No Resection - Refused Treatment
NoResMagnitude	Yes/No	No Resection - Magnitude Not Worth Benefits
NoResCeliacInvolve	Yes/No	No Resection - Celiac Trunk Involvement
NoResSMAInvolve	Yes/No	No Resection - SMA Involvement
NoResHepaticInvolve	Yes/No	No Resection - Hepatic Involvement
NoResIVCInvolve	Yes/No	No Resection - Inferior Vena Cava Involvement
NoResSMVInvolve	Yes/No	No Resection - SMV Involvement
NoResPVInvolve	Yes/No	No Resection - Portal Vein Involvement
NoResCirrhosis	Yes/No	No Resection - Cirrhosis
NoResMetastatic	Yes/No	No Resection - Metastatic

Figure 22: Pancreatic Cancer No Resection Table Schema

Microsoft Acc	ss - [Pan_7_	Path : Fo	rm]																			×
🗃 <u>F</u> ile <u>E</u> dit	<u>V</u> iew <u>I</u> nsert	Format	Re	cords	Iool	s ⊻	∕indov	∾ <u>H</u>	lelp	Ado	be Pl	DF		Туре	a qu	estio	n for	help	•	-	8	×
🛃 • 🗟	🖨 🖪 🖤	1 X B	R	10		≜ ↓	Z.	T.	ð	7	鈉)*	19K		1) {	•	?	Ļ			
	MS Sans Ser	if	÷	10	-	в	I	U		. :	-	٨	- 1	•	-1			•	- •			
- inal Tum	or Histo	loav													-			-				1
					13																	
			5.75	-																		
(from best of imag	ng, FNA, patho	ology, etc)	J																		
)	-																		
)																			
(from best of imag Pathology Fumor Size (cm):)																			
Pathology		able))		м [₹ F	ł: [P		<u> </u>	1									
Pathology Fumor Size (cm):		able))		м [J ₽	ь Г			ž	I									
Pathology Fumor Size (cm):		able))		м [J ₽	n [<u>I*</u>	[

Figure 23: Pancreatic Cancer Pathology Form

Field Name	Data Type Description						
MR	AutoNumber	ID					
MR	Text	Meditech Medical Record Number for Patient					
Histology	Number	Histology					
ResPathT	Number	Resection - Pathology Staging - T					
ResPathN	Number	Resection - Pathology Staging - N					
ResPathM	Number	Resection - Pathology Staging - M					
ResPathR	Number	Resection - Pathology Staging - R					
ResPathV	Number	Resection - Pathology Staging - R					
ResPathSizeX	Number	Resection - Pathology Tumor Size (cm) - Width					
ResPathSizeY	Number	Resection - Pathology Tumor Size (cm) - Height					

Figure 24: Pancreatic Cancer Pathology Table Schema

Microsoft Access - [Follow-u	p Information]	
🗐 File Edit View Insert	Format Records Tools Window Help Adobe PDF Type a question for help.	×
😰 • 🖬 🖬 🚳 🖪 🖤 🛛	¾ ☜ ጫ ⋈ 🤹 🛃 Ў ☜ 🚧 * 🗰 🖬 • Q . –	
Pancreatic Tumor		
Follow-up Informat	tion Add Record Find Record Delete Record 📭	
MR 000726070	Follow-up Window: 9M -	
Date of Visit: 05/11/20	003 Weight (pounds): QOL score:	
	10 @ 1 @ 2 @ 3 @ 4	
Lab Value:		
	umin Alkaline Phosphotase	
CA19-9: Tot	al Bilirubin	
Redeveloped Symptoms:		
🗖 Weight loss	🗖 Biliary colic 🛛 🗖 Pruritis 🗖 Back pain	
how much (pounds)	🗌 🗖 Nausea 👘 Abdominal pain 🗖 Indigestion	
🗖 Jaundice	🗖 Vomiting 👘 Other Specify:	
🗖 Cholecystitis	Clay colored stool	
🗖 Cholangitis	Fatigue	
Status:		
Died Death Date:	5 /11 / 2003	
C NED.		
C A.W.D., Method of Detection	: 🗖 Lab 🔲 Radiologic Evidence 🔲 Clinical Evidence	
Form View		1

Figure 25: Pancreatic Cancer Follow-Up Form

12	Field Name Data Type		Description
I	D	AutoNumber	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
P	1R	Text	Meditech Medical Record Number for Patient
F	UWin	Number	Follow-Up Windows
1	/isitDate	Date/Time	Visit Date
1	Veight	Number	Weight (lbs.)
	OLscore	Number	QoL Score (0-100)
E	COG	Number	ECOG Score (0-4)
L	abCEA	Number	Laboratory - CEA
L	abCA19-9	Number	Laboratory - CA19-9
L	abAlb	Number	Laboratory - Albumin
L	abBili	Number	Laboratory - Bilirubin
L	abAlka	Number	Laboratory - Alkaline phosphotase
19	5xWtloss	Yes/No	Symptoms - Weight Loss
19	5xWtlossP	Number	Symptoms - Weight Loss (lbs.)
19	5xJaun	Yes/No	Symptoms - Jaundice
19	ixChole	Yes/No	Symptoms - Cholecystitis
19	5xChola	Yes/No	Symptoms - Cholangitis
9	5xBC	Yes/No	Symptoms - Biliary Colic
19	5xNau	Yes/No	Symptoms - Nausea
19	5xVom	Yes/No	Symptoms - Vomiting
19	5xCC5	Yes/No	Symptoms - Clay Colored Stool
19	SxFati	Yes/No	Symptoms - Fatigue
19	5xPru	Yes/No	Symptoms - Pruritis
19	ixInd	Yes/No	Symptoms - Indigestion
19	5xAbd	Yes/No	Symptoms - Abdominal Pain
9	5xBack	Yes/No	Symptoms - Back Pain
19	5xOT	Yes/No	Symptoms - Other
19	SXOTSpe	Text	Symptoms - Other - Specify
19	itatus	Number	Status (NED, AWD, Died)
10	DeathDate	Date/Time	Death Date
19	itatusAWDLab	Yes/No	AWD - Lab Evidence
	itatusAWDRad	Yes/No	AWD - Radiology Evidence
	itatusAWDCli	Yes/No	AWD - Clinical Evidence

Figure 26: Pancreatic Cancer Follow-Up Table Schema

3.1.2 Hepatocellular Cancer

L • 🖬 🗞 🎒 🖪 🖪 🖤 🗌	x 🖻 🛍 🗠 🍓 🎒	XI V V V	• • • • 🖻 🗗] /
MS Sans Serif	the second s		2 · A · 1	
recumptive Disers	- I - I - I	the set of the second		
resumptive Diagno	osis al Onsel of	Care		
	-			
resentation				
Data at Fusikation - FCOC I	Desferences Chabus	(1.1.4.C.) M.L.	L. ((L.)	
		Height (in.) Weig	<u>ht (lbs.)</u>	
© 0 0		Height (in.) Weig	<u>ht (lbs.)</u>	
Symptoms	81 62 63 64			
Symptoms Weight Loss	© 1 @ 2 @ 3 @ 4 Biliary Colic	Furitis	Indigestion	
Symptoms Weight Loss How Much (pounds):	© 1 © 2 © 3 © 4 I	Fruritis Abdominal Pain	I™ Indigestion I™ Dysphagia	
Symptoms Weight Loss How Much (pounds): Jaundice	◎ 1 ◎ 2 ◎ 3 ◎ 4 「悪 Biliary Colic 「悪 Nausea 「悪 Vomiting	Fruritis Abdominal Pain Back Pain	Indigestion	
Symptoms Weight Loss How Much (pounds):	© 1 © 2 © 3 © 4 I	Fruritis Abdominal Pain Back Pain	I™ Indigestion I™ Dysphagia	

Figure 27: Hepatocellular Cancer Presentation Form

Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)	
DemEvalDate	Date/Time	Demographics - Date Evaluated by Surgical Oncology	
DemECOG	Number	Demographics - ECOG Score (0-4)	
DemHeight	Number	Demographics - Height in Inches of Patient	
DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission	
SxWtloss	Yes/No	Initial Symptoms - Weight Loss	
SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds	
SxJaun	Yes/No	Initial Symptoms - Juandice	
SxChole	Yes/No	Initial Symptoms - Cholecystitis	
SxChola	Yes/No	Initial Symptoms - Cholangitis	
SxBC	Yes/No	Initial Symptoms - Biliary Colic	
SxNau	Yes/No	Initial Symptoms - Nausea	
SxVom	Yes/No	Initial Symptoms - Vomiting	
SXCCS	Yes/No	Initial Symptoms - Clay Colored Stool	
SxFati	Yes/No	Initial Symptoms - Fatigue	
SxPru	Yes/No	Initial Symptoms - Pruritis	
SxInd	Yes/No	Initial Symptoms - Indigestion	
SxAbd	Yes/No	Initial Symptoms - Abdominal Pain	
SxBack	Yes/No	Initial Symptoms - Back Pain	
SxDyspha	Yes/No	Initial Symptoms - Dysphagia	
SxSatiety	Yes/No	Initial Symptoms - Early Satiety	
SXOT	Yes/No	Initial Symptoms - Other	
SxOTSpe	Text	Initial Symptoms - Other - Specify	

Figure 28: Hepatocellular Cancer Presentation Table Schema

Microsoft Access - [HCC_2]	_History : Form]			
🗐 Eile Edit View Insert	Format Records Iools y	∦indow <u>H</u> elp Ado <u>b</u> e PDF	Type a question for help	8 ×
🔟 • 🖬 🗞 🥌 🗛 🖤	🐰 🖻 💼 🗠 🍓 🛃	XI V T V M >*	w 🖻 🗗 🔄 🖸 .	
· · · · · · · · · · · · · · · · · · ·	• B	1 U 🗏 🗏 🗏 🖄	- <u>A</u> - <u>2</u> - <u>-</u> -	• .
Medical History				
Comorbidities	-	Cancer History		
Heart Failure Ischemic Heart Disease Respiratory Renal Failure Hypertension Bleeding Disorder Social History	 Mainutrition Liver Failure/Cirrhosis Diabetes Less than Six Months Greater than Six Months Oral Agents Diet Control 	Patient Prior Dx: Chemo I Radiation Father Dx: Other Relation: Related Dx:	Surgery	
	ar Drug Use rmental Exposure	Other Relation: Related Dx:		
Form View				

Figure 29: Hepatocellular Cancer Medical History Form

Field Name Data Type		Description							
)D	AutoNumber	ID							
MR	Text	Meditech Medical Record Number for Patient							
CxHF	Yes/No	Comorbidities - Heart Failure							
CxIHD									
CxResp	Yes/No	Comorbidities - Respiratory							
CxDiab	Yes/No	Comorbidities - Diabetes							
CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral							
CxDiabDiet	Yes/No	Comorbidities - Diabetes - Insulin - Diet Control							
CxDiabOnset	Number	Comorbidities - Diabetes - Onset (1=Less than six months, 2 =Greater than six months)							
CxRF	Yes/No	Comorbidities - Renal Failure							
CxHyper	Yes/No	Comorbidities - Hypertension							
CxBleed	Yes/No	Comorbidities - Bleeding Disorder							
CxLiver	Yes/No	Comorbidities - Liver Failure							
CxMal	Yes/No	Comorbidities - Malnutrition							
CxPriorCancer	Number	Comorbidities - Prior Cancer Dx							
CxPriorCancerChemo	Yes/No	Comorbidities - Prior Cancer Dx - Chemo							
CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation							
CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery							
SHCigarette	Yes/No	Social History - Cigarettes (significant use)							
SHAlcohol	Yes/No	Social History - Alcohol (significant use)							
SHDrugUse	Yes/No	Social History - Drug Use							
SHExposure	Yes/No	Social History - Environmental Exposure							
SHOther	Yes/No	Social History - Other							
SHOtherS	Text	Social History - Other - Specify							
FamilyFatherDx	Number	Family History - Father Dx							
FamilyMotherDx	Number	Family History - Mother Dx							
FamilyOther1	Text	Family History - Other1							
FamilyOther1Dx	Number	Family History - Other1 Dx							
FamilyOther2	Text	Family History - Other2							
FamilyOther2Dx	Number	Family History - Other2 Dx							

Figure 30: Hepatocellular Cancer Medical History Table Schema

	7	t <u>R</u> e	cords	Tool	1995	ndow	Help	Adobe	PDF	Туре	e a que	stion for l	help	
🔛 • 🔛 📆	🖨 🖪 🖤 🔏 🗉		160	6	₹↓ Z	1 3	百	V	h >+ 1%			*	2.	
	MS Sans Serif	÷	8	4	в	ΙŪ	F	王 王	2 - 1	A -	1 -	-		
CEA:	Albumin	-		AL	.K	1		ALT			1	14		
CA19-9:	Total Bilirubin			AS		-		ALI			8			

Figure 31: Hepatocellular Cancer Serum Studies Form

Field Name	Data Type Description						
MR	AutoNumber	ID					
MR	Text	Meditech Medical Record Number for Patient					
LabCEA	Number	Laboratory - CEA					
LabCA19-9	Number	Laboratory - CA19-9					
LabAlb	Number	Laboratory - Albumin					
LabBili	Number	Laboratory - Bilirubin					
LabAlka	Number	Laboratory - Alkaline phosphotase					
LabALT	Number	Laboratory - ALT					
LabAST	Number	Laboratory - AST					
LabAmylase	Number	Laboratory - Amylase					
S 28							

Figure 32: Hepatocellular Cancer Serum Studies Table Schema

Microsoft Access - [HCC_3b_Diag : Form]
🗃 File Edit View Insert Format Records Iools Window Help Adobe PDF Type a question for help 🗸 🗗
·⊻·■┓⊜Q♥☆®高い@針科孯┓▽ぬ艸咪◎@@┪·Q.
MS Sans Serif • 8 • B I ∐ ≣ ≣ ≣ ⊉ • ▲ • ⊿ • □ • .
Diagnostic Studies
Childs Class:
Primary Tumor Staging: T 💽 N 💽 R: 🔽 Grade:
<u>CT Scan</u>
Date: 🗾 # of Tumors: 💽
₩ Bilobar Size of largest tumor (cm);
🗑 Evidence of metastatic disease 🖉 Portal Vein 🕅 Hepatic Vein 🕅 IVC
III Portal HTN
I CXR Mets
🖩 Prior Chemotherapy Type: 💽 Start Date: End Date:
🕅 Avastin 🕼 Leukovorin 🕼 Fluorouracil (5-FU) 📓 Irinotecan 🕼 Other, specify:
🕅 Capecitabine 🕅 Levamasole 🕅 FUDR 🕅 Oxaliplatin
🖩 Erbitux ា Mitomycin 🕅 Gemcitabine 🕅 Taxol
J Form View

Figure 33: Hepatocellular Cancer Diagnostic Imaging Form

-	Field Name	Data Type	Description
	1D	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
	ChildsC	Number	Childs Class
	TStage	Number	Imaging T-Stage
	NStage	Number	Imaging N-Stage
	RStage	Number	Imaging R-Stage
	CTDate	Date/Time	CT Date
	CTCount	Number	CT Tumor Count
	Bilobar	Yes/No	Bilobar
	TumorSize	Number	Size of Largest Tumor
	Invasion PV	Yes/No	Invasion - Portal Vein
	Invasion HV	Yes/No	Invasion - Hepatic Vein
	Invasion IVC	Yes/No	Invasion - Inferior Vena Cava
	Metastatic	Yes/No	Metastatic Evidence
	Portal	Yes/No	Portal HTN
	CXR	Yes/No	Chest X-Ray
	CXRMets	Yes/No	Chest X-Ray - Metastatic Evidence
	PChemo	Yes/No	Prior Chemo
	PChemoType	Number	Prior Chemo - Type
	PChemoStartD	Date/Time	Prior Chemo - Start Date
	PChemoEndD	Date/Time	Prior Chemo - End Date
1	PChemo AVA	Yes/No	Prior Chemo - Avastin
	PChemo Cap	Yes/No	Prior Chemo - Capecitabine
1	PChemo Erb	Yes/No	Prior Chemo - Erbitux
	PChemo Flu	Yes/No	Prior Chemo - Fluorouracil (5-FU)
1	PChemo FUDR	Yes/No	Prior Chemo - FUDR
	PChemo Gem	Yes/No	Prior Chemo - Gemcitabine
1	PChemo Iri	Yes/No	Prior Chemo - Irinotecan
	PChemo Leu	Yes/No	Prior Chemo - Leukovorin
1	PChemo Lev	Yes/No	Prior Chemo - Levamasole
	PChemo Mit	Yes/No	Prior Chemo - Mitomycin
	PChemo Oxa	Yes/No	Prior Chemo - Oxaliplatin
	PChemo Tax	Yes/No	Prior Chemo - Taxol
	PChemo Oth	Yes/No	Prior Chemo - Other
	PChemo OS	Text	Prior Chemo - Other - Specify

Figure 34: Hepatocellular Cancer Diagnostic Imaging Table Schema

Microsoft Access	- [HCC_4_Prelim : Form]
🖪 File Edit View	/ Insert Format Records Tools Window Help Adobe PDF Type a question for help 🔽 🗗 🗙
	Q.♥ % 6 2 ¾ № Ø Ø Ø 0
Disease Stag	<u>16</u>
 B or C cirrhosis Non-cirrhotic or Chil 	Correction Contract Model Contract Contrect Contract Contrect Contrect Contrect Contract Contract Contract
Pre-Surgical	Outlook
Potentially Rese	ectable
Cocally Advance	ed/Unresectable
Metastatic or Education	quivocal Findings
Pursue Rese Tumor Volume:	Resectable: Got to Laparoscopy:
Form View	

Figure 35: Hepatocellular Cancer Preliminary Outlook Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
Cirr	Number	Cirrhosis Type
CirrBC	Number	Cirrhosis BC Type
TranTime	Number	Transplant Time
TranTime2	Number	Transplant Time > 2 Months
PreOutlook	Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic or Equivocal Findings)
TumorVol	Number	Tumor Volume
SurgTreat	Number	Initial Surgical Decision (Resect, Ablate, Resect/Ablate)

Figure 36: Hepatocellular Cancer Preliminary Outlook Table Schema

Picrosoft Access - [H	ICC_5_Treatment : Form]			-O×
E Eile Edit View	Insert Format Records	<u>T</u> ools <u>W</u> indow <u>H</u> elp Ac	do <u>b</u> e PDF	Type a question for help	8 ×
1 🛛 🖌 🖬 📾 🖾	à 🖤 👗 🖻 💼 🗠		1 1 × 100	P 6 . Q.	
*	*	• B I U =	≡ 2 • /	<u>\</u> - <u>/</u> -	• .
Treatment Cou	<u>ırse</u>				
Resection		Staging Laparos	copy/Lapa	rotomy	
Radiation		🗏 Palliative Measu	ires		
Adjuvant Neoad Neoa	djuvant 🖗 Both	🕅 Bypass	HAL		
Chemotherapy		🕅 Gastrostomy Tube	🕅 PV Shunt	s	
Adjuvant S Neoad	fjuvant 🖗 Both	🕅 Jejunstomy Tube	🕅 Pall. Steni	ting	
🕅 Avastin	🕅 Leukovorin	🕅 Celiac Block	🕅 Pall. Radi	ation	
🕅 Capecitabine	🕅 Levamasole	🕅 Paracentesis	🕅 Pall. Rese	ection	
🕅 Erbitux	🕅 Mitomycin	Thoracentesis	🕅 Other - Sp	becify:	
🕅 Fluorouracil (5-FU)	🕅 Oxaliplatin	Transfusion			
FUDR	🕅 Taxol			and the second second	
👿 Gemcitabine	🕅 Other - Specify:	Experimental Pro	otocol (vaco	cine, etc.)	
linotecan		🗏 <u>Genetic Counsel</u>	ling		
3=0.					
Form View					

Figure 37: Hepatocellular Cancer Treatment Form

	Field Name	Data Type	Description
? •	JD	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
	TxResect	Yes/No	Treatment - Resection
1121	TxLap	Yes/No	Treatment - Laparoscopy
	TxRadia	Yes/No	Treatment - Radiation
1	TXRadiaAdju	Number	Treatment - Radiation - Adjuvancy
-	TxChemo	Yes/No	Treatment - Chemo
	TXChemoAdju	Number	Treatment - Chemo - Adjuvancy
	TxChemoAVA	Yes/No	Treatment - Chemo - Avastin
1	TxChemoCap	Yes/No	Treatment - Chemo - Capecitabine
-	TxChemoErb	Yes/No	Treatment - Chemo - Erbitux
	TxChemoFlu	Yes/No	Treatment - Chemo - Fluorouracii (5-FU)
	TxChemoFUDR	Yes/No	Treatment - Chemo - FUDR
1	TxChemoGem	Yes/No	Treatment - Chemo - Gemcitabine
-	TxChemoIri	Yes/No	Treatment - Chemo - Irinotecan
	TxChemoLeu	Yes/No	Treatment - Chemo - Leukovorin
	TxChemoLev	Yes/No	Treatment - Chemo - Levamasole
1	TxChemoMit	Yes/No	Treatment - Chemo - Mitomycin
	TxChemoOxa	Yes/No	Treatment - Chemo - Oxaliplatin
	TxChemoTax	Yes/No	Treatment - Chemo - Taxo
	TxChemoOth	Yes/No	Treatment - Chemo - Other
1	TxChemoOS	Text	Treatment - Chemo - Other - Specify
	TxPal	Yes/No	Treatment - Palliation
	TxPalRes	Yes/No	Treatment - Palliation - Pall, Resection
	TxPalBypass	Yes/No	Treatment - Palliation - Bypass
	TxPalCeliac	Yes/No	Treatment - Palliation - Celiac Block
-	TxPalPara	Yes/No	Treatment - Pallation - Paracentesis
	TxPalTho	Yes/No	Treatment - Pallation - Thoracentesis
	TxPalRad	Yes/No	Treatment - Palliation - Pall, Badiation
_	TxPalTrans	Yes/No	Treatment - Pallation - Transfusion
	TxPalStens	Yes/No	Treatment - Palliation - Pall. Stenting
	TxPalPV	Yes/No	Treatment - Pallation - PV Shunts
		Yes/No	Treatment - Pallation - HAI
-	TxPalGasTube	Yes/No	Treatment - Pallation - Gastrostomy Tube
	TxPalJeiTube	Yes/No	Treatment - Pallation - Jejunstomy Tube
	TxPalOth	Yes/No	Treatment - Palliation - Other
	TxPalOS	Text	Treatment - Pallation - Other - Specify
-	TXEXp	Yes/No	Treatment - Experimental protocol (ie. vaccine)
	TxGene	Yes/No	Treatment - Gene Counseling
	TXGCHE	respino	I reacher to the course ing

Figure 38: Hepatocellular Cancer Treatment Table Schema

	ormat Records Iools Window Help Adobe PDF Type a question for help 👻	- 8 ×
¥• ₽ ¶ 60 ,♥ 3 -	6 ₽ € ▷ 6 1 1 1 ỹ 7 ∨ A >• K ☎ ⊡ ⊕ • Q. • B 1 ⊻ ≡ ≡ Ѯ• ▲• <u>/</u> • □• .	Į
blation		-
All Metastatic Disease ablated Clear Gross Margins (Eye, U/S) CT Guidance Ultrasound Guidance Hepatoduodenal LN Biopsied Hepatoduodenal LN Clear	# of Mets Ablated: Largest Met Ablated (cm): Locations of Ablated Lesions Seg Locations of Largest Ablated Lesions Seg Number of Ablated Lesions	_
		Ŀ

Figure 39: Hepatocellular Cancer Ablation Form

Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
Abl_MetsAll	Yes/No	Ablation - All Mets Ablated	
Abl_MetsClear	Yes/No	Ablation - Clear Gross Margins	
Abl_MetsCount	Number	Ablation - Number of Mets	
Abl_MetsLargest	Number	Ablation - Largest Met Ablated (cm)	
Abl_CT	Yes/No	Ablation - CT Guidance	
Abl_US	Yes/No	Ablation - US Guidance	
Abl_LNBio	Yes/No	Ablation - Hepatoduodenal LN Biopsed	
Abl_LNClear	Yes/No	Ablation - Hepatoduodenal LN Clear	
Abl_LesionSeg	Number	Ablation - Locations of Ablated Lesions	
Abl_LesionCount	Number	Ablation - Number of Ablated Lesions	
Abl_LesionLargest	Number	Ablation - Locations of Largest Ablated Lesions	

Figure 40: Hepatocellular Cancer Ablation Table Schema

Microsoft Access - [HCC_7a_Res : Form]	١×
🖼 File Edit View Insert Format Records Iools Window Help Adobe PDF Type a question for help 👻 _ r	9 ×
K·□ 0 0 0 0 × x 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
- MS Sans Serif • 8 • B I U = = = 🖄 • 🗛 • 🖉 • 🔽 •	
f Resection is Performed	
Date of Admission: Date of Surgery: OR Time (hr):	
Procedure Type 🔽 Days in ICU:	
CVP cm/H20 During Hepatic Transection/Puncture: cm/h20:	
Tidal Volume During Hepatic Transection/Puncture (cc):	
Pathological Staging: T 🔽 N 💌 M 💌 R:	
Other Organs Resected: Estimated Blood Loss (cc):	
🕅 Transfusion If yes, units: 🚺 Methods: 🕅 FFP 🕅 Cell Saver	
🗑 Tissue Link 🖉 Finger Fracture 🕼 Argon Beam 🖉 Staplers for Structures 🗐 RFA	
🗰 CVSA 🗰 Clamp/Crush 🗰 Staplers for Parynchyma 🗰 Fibrin Glue 🗰 Cryoablation	
₩ Wound Infection Congruent with Post Op Care Path M NG/Gastrostomy Drainage > 7 Days	
Ecological in work of the congress in the second of t	
Example at the second s	
🗰 Liver Insufficiency (Total Bilirubin >5 Inpatient) If Yes, Total Bilirubin:	
Discharge Status: Date of Discharge/Death:	
Metastatic gastric ademo with resectable primary tumor:	
🖉 Primary Tumor Grossly Symptomatic Volume of Metastatic Disease: 🗾 🗾	
Form View	

Figure 41: Hepatocellular Cancer Resection Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
DAdm	Date/Time	Date of Admission
DSurg	Date/Time	Date of Surgery
ORTime	Number	OR Time
PxType	Number	Procedure Type
Organs	Text	Other Organs Resected
BloodLoss	Number	Blood Loss (cc)
CVP	Yes/No	CVP cm/H20 During Hepatic Transection/Puncture
CVPcm	Number	CVP cm/H20
Tidalcc	Number	Tidal Volume During Hepatic Transection/Puncture (cc)
Transfusion	Yes/No	Transfusion - Needed?
T_Units	Number	Transfusion - Units
T_FFP	Yes/No	Transfusion - Fresh Frozen Plasma
T Cell	Yes/No	Transfusion - Cell Saver
T Tissue	Yes/No	Transfusion - Tissue Link
T CVSA	Yes/No	Transfusion - CVSA
T Finger	Yes/No	Transfusion - Finger Fracture
r Clamp	Yes/No	Transfusion - Clamp/Crush
T Argon	Yes/No	Transfusion - Argon Beam
T Pary	Yes/No	Transfusion - Staplers for Parynchyma
T Struct	Yes/No	Transfusion - Staplers for Structures
TGlue	Yes/No	Transfusion - Fibrin Glue
T RFA	Yes/No	Transfusion - RFA
T_Cry	Yes/No	Transfusion - Cryoablation
CUdays	Number	Days in ICU
Infection	Yes/No	PO - Wound Infection
eak	Yes/No	PO - Leak
Congruent	Yes/No	PO - Congruent Post-Op Path
Divergent	Yes/No	PO - Divergent Post-Op Path
NG	Yes/No	PO - NG/Gastrostomy Drainage > 7 Days
Abdominal	Yes/No	PO - Abdominal Collection
PulmCx	Yes/No	PO - Pulmonary Complications
Catheter	Yes/No	PO - Catheter Infections
Drains	Yes/No	PO - Drains
RenalInsuf	Yes/No	PO - Renal Insufficiency
LiverInsuf	Yes/No	PO - Liver Insufficiency
LI TB	Number	PO - Liver Insufficiency - Total Bilirubin
DDischarge	Date/Time	Date of Discharge
DischStatus	Number	Discharge Status
DDeath	Date/Time	Date of Death

Figure 42: Hepatocellular Cancer Resection Table Schema



Figure 43: Hepatocellular Cancer No Resection Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
NoResEvalDate	Date/Time	No Resection - Decision Date
NoResNoHandle	Yes/No	No Resection - Couldn't Handle Proposed Treatment
NoResRefused	Yes/No	No Resection - Refused Treatment
NoResMagnitude	Yes/No	No Resection - Magnitude Not Worth Benefits
NoResCeliacInvolve	Yes/No	No Resection - Celiac Trunk Involvement
NoResSMAInvolve	Yes/No	No Resection - SMA Involvement
NoResHepaticInvolve	Yes/No	No Resection - Hepatic Involvement
NoResIVCInvolve	Yes/No	No Resection - Inferior Vena Cava Involvement
NoResSMVInvolve	Yes/No	No Resection - SMV Involvement
NoResPVInvolve	Yes/No	No Resection - Portal Vein Involvement
NoResCirrhosis	Yes/No	No Resection - Cirrhosis
NoResMetastatic	Yes/No	No Resection - Metastatic

Figure 44: Hepatocellular Cancer No Resection Table Schema

🗃 File Edit View Insert	Format	Rec	ords	Tools	s <u>W</u> ir	ndow	Help	Ad	o <u>b</u> e P	DF		Гуре	a que	stion	for h	elp.	-	_ é	, s
¥• 🖬 🗞 🍯 🖪 🖤	X @	B	6			ļŢ	百	∇	44	•*	X	P	Ð	1	- [2.			
MS Sans Serif		-	10	-	в.	τ <u>u</u>	F	Ŧ	-	۵	A	•	1	•	`]•		• .		
inal Tumor Histol	vp											-							
			-																
from best of imaging, FNA, patholo	gy, etc)		-																
(from best of imaging, FNA, patholo	gy, etc)	2 2																	
		<u></u>																	
athology (if availa	<u>ble)</u>	<u>1</u>																	
Pathology (if availa	ble)			мГ			B,			-									
(from best of imaging, FNA, patholo Pathology (if availa Fumor Size (cm):	<u>ble)</u>			м [R:												
Pathology (if availa	ble)			мГ			R:			Ī									

Figure 45: Hepatocellular Cancer Pathology Form

Field Name	Data Type	Description	
MR	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
Histology	Number	Histology	
ResPathT	Number	Resection - Pathology Staging - T	
ResPathN	Number	Resection - Pathology Staging - N	
ResPathM	Number	Resection - Pathology Staging - M	
ResPathR	Number	Resection - Pathology Staging - R	
ResPathV	Number	Resection - Pathology Staging - R	
ResPathSizeX	Number	Resection - Pathology Tumor Size (cm) - Width	
ResPathSizeY	Number	Resection - Pathology Tumor Size (cm) - Height	

Figure 46: Hepatocellular Cancer Pathology Table Schema

	rmat <u>R</u> ecords <u>T</u> ools <u>W</u> ind		Type a question for help	
NEW LARGE STATE AND A STATE OF A DATE	the share in the second s			
Hepatocellular Can	icer			
Follow-up Informat	tion Add Record	Find Record Delete Reco	rd 📭	
MR	Follow-up Window:	•		
Date of Visit:	Weight (pounds):	QOL score:	r1	
Cascing assessed in the second s	0 @ 1 @ 2 @ 3	@ 4	4	
Content a Restaurant and Analysis (C	0 61 62 63			
Lab Value:	·			
	SOS	Ikaline phosphotase		
CA19-9: Tota	al Bilirubin S	GOT #Name? SGPT	#Name?	
Redeveloped symptoms:				
🗖 Weight loss	🔲 Biliary colic	🗖 Pruritis 🛛 🗖	ndigestion	
how much (pounds)	🗌 🗖 Nausea	🗖 Abdominal pain		
🛛 🗖 Jaundice	🗖 Vomiting	🗖 Back pain		
Cholecystitis	Clay colored stool	C Other Specify:		
🗖 Cholangitis	Fatigue			
Status:				
C Died Death Date:			-	
C N.E.D.				
C A.W.D., method of detection:	🗖 Lab 🔲 Radiologic	evidence 🛛 🗖 Clinical evidenc	e	
		en e		
Record: 14 4 1 1 1 1	▶* of 1			

Figure 47: Hepatocellular Cancer Follow-Up Form

Field Name		Data Type	Description
	ID	AutoNumber	· · · · · · · · · · · · · · · · · · ·
	MR	Text	Meditech Medical Record Number for Patient
1	FUWin	Number	Follow-Up Windows
1	VisitDate	Date/Time	Visit Date
	Weight	Number	Weight (lbs.)
	QOLscore	Number	QoL Score (0-100)
	ECOG	Number	ECOG Score (0-4)
1	LabCEA	Number	Laboratory - CEA
	LabCA19-9	Number	Laboratory - CA19-9
1	LabAlb	Number	Laboratory - Albumin
1	LabBili	Number	Laboratory - Bilirubin
1	LabAlka	Number	Laboratory - Alkaline phosphotase
	SxWtloss	Yes/No	Symptoms - Weight Loss
	SxWtlossP	Number	Symptoms - Weight Loss (lbs.)
	SxJaun	Yes/No	Symptoms - Jaundice
	SxChole	Yes/No	Symptoms - Cholecystitis
1	SxChola	Yes/No	Symptoms - Cholangitis
	SxBC	Yes/No	Symptoms - Biliary Colic
	SxNau	Yes/No	Symptoms - Nausea
1	SxVom	Yes/No	Symptoms - Vomiting
	SxCCS	Yes/No	Symptoms - Clay Colored Stool
	SxFati	Yes/No	Symptoms - Fatigue
	SxPru	Yes/No	Symptoms - Pruritis
	SxInd	Yes/No	Symptoms - Indigestion
1	SxAbd	Yes/No	Symptoms - Abdominal Pain
	SxBack	Yes/No	Symptoms - Back Pain
	SxOT	Yes/No	Symptoms - Other
	SxOTSpe	Text	Symptoms - Other - Specify
1	Status	Number	Status (NED, AWD, Died)
1	DeathDate	Date/Time	Death Date
1	StatusAWDLab	Yes/No	AWD - Lab Evidence
	StatusAWDRad	Yes/No	AWD - Radiology Evidence
	StatusAWDCli	Yes/No	AWD - Clinical Evidence

Figure 48: Hepatocellular Cancer Follow-Up Table Schema

3.1.3 Gall Bladder/Biliary Cancer

	• 8 • B I	⊻ ≡ ≡ ≡ ≤	• ₩ @ @ • Q. 2 • <u>A</u> • <u>∠</u> • □ • □	
		17	≥ • <u>▲</u> • <u>⊿</u> • <mark>-</mark> • -	• .
umptive Diagnosis a	t Onset of Ca	<u>e</u>		
entation				
			i	
of Evaluation ECOG Performa	Contraction of the second s	ht (in.) Weight (I	<u>bs.</u>]	
a second case	2 @ 3 @ 4			
otoms			- 1 - P - P	
			Indigestion	
	/		🛙 Dysphagia	
aundice 🕅	Vomiting 🕅	Back Pain 🕅	Early Satiety	
holecystitis 🕅	Clay Colored Stool	Other Specify:		
and a second	Fatigue			
holangitis 🕅	raugue			

Figure 49: Gall Bladder/Biliary Cancer Presentation Form

Field Name Data Ty		Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)	
DemEvalDate	Date/Time	Demographics - Date Evaluated by Surgical Oncology	
DemECOG	Number	Demographics - ECOG Score (0-4)	
DemHeight	Number	Demographics - Height in Inches of Patient	
DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission	
SxWtloss	Yes/No	Initial Symptoms - Weight Loss	
SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds	
SxJaun	Yes/No	Initial Symptoms - Juandice	
SxChole	Yes/No	Initial Symptoms - Cholecystitis	
SxChola	Yes/No	Initial Symptoms - Cholangitis	
SxBC	Yes/No	Initial Symptoms - Biliary Colic	
SxNau	Yes/No	Initial Symptoms - Nausea	
SxVom	Yes/No	Initial Symptoms - Vomiting	
SXCCS	Yes/No	Initial Symptoms - Clay Colored Stool	
SxFati	Yes/No	Initial Symptoms - Fatigue	
SxPru	Yes/No	Initial Symptoms - Pruritis	
SxInd	Yes/No	Initial Symptoms - Indigestion	
SxAbd	Yes/No	Initial Symptoms - Abdominal Pain	
SxBack	Yes/No	Initial Symptoms - Back Pain	
SxDyspha	Yes/No	Initial Symptoms - Dysphagia	
SxSatiety	Yes/No	Initial Symptoms - Early Satiety	
SXOT	Yes/No	Initial Symptoms - Other	
SxOTSpe	Text	Initial Symptoms - Other - Specify	

Figure 50: Gall Bladder/Biliary Cancer Presentation Table Schema

	Window Help Adobe PDF Type a question for help ▼ _ 5
Medical History	
Comorbidities Image: Heart Failure Image: Malnutrition Image: Ischemic Heart Disease Image: Liver Failure/Cirrhosis Image: Ischemic Heart Disease Image: Liver Failure/Cirrhosis Image: Respiratory Image: Diabetes Image: Renal Failure Image: Liver Failure/Cirrhosis Image: Renal Failure Image: Cirrhosis Image: Renal Failure <td>Cancer History Patient Prior Dx: Im Chemo Im Chemo Im Adiation Im Chemo Im Adiation Im Chemo Im Adiation Im Chemo Im Adiation <</td>	Cancer History Patient Prior Dx: Im Chemo Im Chemo Im Adiation Im Chemo Im Adiation Im Chemo Im Adiation Im Chemo Im Adiation <
Alcohol Use Environmental Exposure Other - Specify:	Related Dx:

Figure 51: Gall Bladder/Biliary Cancer Medical History Form

Field Name Data Typ		Description
1D	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
CxHF	Yes/No	Comorbidities - Heart Failure
CxIHD	Yes/No	Comorbidities - Ischemic Heart Disease
CxResp	Yes/No	Comorbidities - Respiratory
CxDiab	Yes/No	Comorbidities - Diabetes
CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral
CxDiabDiet	Yes/No	Comorbidities - Diabetes - Insulin - Diet Control
CxDiabOnset	Number	Comorbidities - Diabetes - Onset (1=Less than six months, 2 =Greater than six months)
CxRF	Yes/No	Comorbidities - Renal Failure
CxHyper	Yes/No	Comorbidities - Hypertension
CxBleed	Yes/No	Comorbidities - Bleeding Disorder
CxLiver	Yes/No	Comorbidities - Liver Failure
CxMal	Yes/No	Comorbidities - Malnutrition
CxPriorCancer	Number	Comorbidities - Prior Cancer Dx
CxPriorCancerChemo	Yes/No	Comorbidities - Prior Cancer Dx - Chemo
CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation
CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery
SHCigarette	Yes/No	Social History - Cigarettes (significant use)
SHAlcohol	Yes/No	Social History - Alcohol (significant use)
SHDrugUse	Yes/No	Social History - Drug Use
SHExposure	Yes/No	Social History - Environmental Exposure
SHOther	Yes/No	Social History - Other
SHOtherS	Text	Social History - Other - Specify
FamilyFatherDx	Number	Family History - Father Dx
FamilyMotherDx	Number	Family History - Mother Dx
FamilyOther1	Text	Family History - Other1
FamilyOther1Dx	Number	Family History - Other1 Dx
FamilyOther2	Text	Family History - Other2
FamilyOther2Dx	Number	Family History - Other2 Dx

Figure 52: Gall Bladder/Biliary Cancer Medical History Table Schema

🔎 Microsoft Acc	ess - [GB_3a_Serum :	Form]				<u>_ ×</u>
🗐 File Edit	View Insert Format	<u>R</u> ecords <u>T</u> o	ols <u>W</u> indow (Help Adobe PDF	Type a question fo	or help 🔽 🗕 🗗 🗙
😥 - 🖬 🗞	6 Q 🖤 🐰 🖻	B 10 8	AL ZI V		x 8 6 6	• 🛛 .
E	MS Sans Serif	• 8 •	BIU	=== 🏖	• <u>A</u> • <u>4</u> • [
Serum Stud	lies					
CEA: CA19-9:	Albumin Total Bilirubin		ILK 🔽	ALT Amylase:		
H. Pylori Status: C	Positive C Negative					
Form View						11

Figure 53: Gall Bladder/Biliary Cancer Serum Studies Form

Carlos and the Constant		
AutoNumber	ID	
Text	Meditech Medical Record Number for Patient	
Number	Laboratory - CEA	
Number	Laboratory - CA19-9	
Number	Laboratory - Albumin	
Number	Laboratory - Bilirubin	
Number	Laboratory - Alkaline phosphotase	
Number	Laboratory - ALT	
Number	Laboratory - AST	
Number	Laboratory - Amylase	
	Number Number Number Number Number Number Number	Number Laboratory - CEA Number Laboratory - CA19-9 Number Laboratory - Albumin Number Laboratory - Bilirubin Number Laboratory - Bilirubin Number Laboratory - Alkaline phosphotase Number Laboratory - ALT Number Laboratory - AST

Figure 54: Gall Bladder/Biliary Cancer Serum Studies Table Schema

	Insert Format Records Iools Window Help Adobe PDF Type a question for help ▼ _ & >
iagnostic Imagi	
	Scan or Ultrasound Duplex
Date of Procedure:	
Status of Main PV:	Copen Cancased Coccluded CAbuts Can't Tell
Right Portal Vein:	© Open @ Encased @ Occluded @ Abuts @ Can't Tell
Left Portal Vein:	Open Encased Cocluded Abuts Can't Tell
Hepatic Arteries:	@ Open @ Encased @ Occluded @ Abuts @ Can't Tell
🕅 Mass Visible	
Extent of Involveme	ent of Hilar Biliary Tree:
🕅 Unknown	🕅 Right Main Duct 🕅 Left Secondary Branch 🕅 Left Tertiary Branch
🕅 Left Main Duct	🗑 Bifurcation 🖉 Right Secondary Branch 🕅 Right Tertiary Branch
Volumetric Assessment	- Proposed Liver Remnant: @ Donecc @ Not Done
Potential Maneuvers to	Increase Safety of Resection: @ Portal Vein Embolization @ Stents for Drainage @ None
Chest X-Ray (CX	
and the second se	
Percutaneous T Date of Procedure:	Transhepatic Cholangiography (PTC)
	≤ I Type: © Internal © External
Stenting	

Figure 55: Gall Bladder/Biliary Cancer Diagnostic Imaging Form

Field Name	Data Type	Description	
)D	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
CXRDx	Yes/No	CXR - Diagnosis	
CTDx	Yes/No	CT - Diagnosis	
CTEvalDate	Date/Time	CT - Date Evaluated	
CTMainPVClass	Number	CT - Main Portal Vein Involvement Class	
CTRightPVClass	Number	CT - Right Portal Vein Involvement Class	
CTLeftPVClass	Number	CT - Left Portal Vein Involvement Class	
CTHepaticClass	Number	CT - Hepatic Arteries Involvement	
CTMassVisible	Yes/No	CT - Mass Visible	
CTHBUnknown	Yes/No	CT - Hilar Biliary Tree Involvement - Unknown	
CTHBLeft	Yes/No	CT - Hilar Biliary Tree Involvement - Left Main Duct	
CTHBRight	Yes/No	CT - Hilar Biliary Tree Involvement - Right Main Duct	
CTHBBifurcation	Yes/No	CT - Hilar Biliary Tree Involvement - Bifurcation	
CTHBLeft2nd	Yes/No	CT - Hilar Biliary Tree Involvement - Left Secondary Duct	
CTHBRight2nd	Yes/No	CT - Hilar Biliary Tree Involvement - Right Secondary Duct	
CTHBLeft3rd	Yes/No	CT - Hilar Biliary Tree Involvement - Left Tertiary Duct	
CTHBRight3rd	Yes/No	CT - Hilar Biliary Tree Involvement - Right Tertiary Duct	
CTVolLiverDone	Number	CT - Volumetric Assessment - Done	
CTVolLiverCC	Number	CT - Volumetric Assessment - CCs	
CTSafetyMan	Number	CT - Potential Manuevers	
PTCDX	Yes/No	PTC - Diagnosis	
PTCEvalDate	Date/Time	PTC - Date Evaluated	
PTCStent	Yes/No	PTC - Stent	
PTCStentType	Number	PTC - Stent Type	

Figure 56: Gall Bladder/Biliary Cancer Diagnostic Imaging Table Schema

Microsoft Access - [GB_4_Prelim : Form]		
🛛 File Edit View Insert Format Recon	ds <u>T</u> ools <u>W</u> indow <u>H</u> elp Ado <u>b</u> e PDF Type	a question for help 🛛 👻 🗕 🗗 🕽
¥• 🛛 🔁 🗿 🗛 🖤 🐰 🖻 🖻 🔺	이 🚳 🛃 💱 🚡 🖓 🏘 🕨 🕷 🖆	0.
*	• B Z U = = = A • A •	⊿ • □ • □ • •
re-Surgical Outlook		
Potentially Resectable		
Locally Advanced/Unresectable		
Metastatic or Equivocal Findings		
	-	

Figure 57: Gall Bladder/Biliary Cancer Preliminary Outlook Form

1194	d Nome 1	Data Type	Description
ĪD	Search	AutoNumber	ID
MR		Text	Meditech Medical Record Number for Patient
PreOutlook		Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic or Equivocal Findings)

Figure 58: Gall Bladder/Biliary Cancer Preliminary Outlook Table Schema

🕄 Eile Edit View	Insert Format Records	8 2 ↓ 3 ↓ 1 y 7 v	obe PDF Type a question for I	help - 8 > ② -
*	•	→ B I U E Ξ	≡ <u>∆</u> · <u>∆</u> · <u>∠</u> ·	• • •
Freatment Cou	irse			
Resection		🖩 Staging Laparos	copy/Laparotomy	
Radiation		🖩 <u>Palliative Measu</u>	res	
Adjuvant Neoad	juvant 🖗 Both	🕅 Bypass	MAL	
Chemotherapy		🕅 Gastrostomy Tube	🕅 PV Shunts	
Adjuvant Neoad	· · · · · · · · · · · · · · · · · · ·	🕅 Jejunstomy Tube	🕅 Pall. Stenting	
🕅 Avastin	🕅 Leukovorin	🕅 Celiac Block	🕅 Pall. Radiation	
🕅 Capecitabine	🕅 Levamasole	🕅 Paracentesis	🕅 Pall. Resection	
🕅 Erbitux	🕅 Mitomycin	🕅 Thoracentesis	🕅 Other - Specify:	
🕅 Fluorouracil (5-FU)	🕅 Oxaliplatin	🕅 Transfusion		
🕅 FUDR	🕅 Taxol	Experimental Pro		
🕅 Gemcitabine	🕅 Other - Specify:		otocol (vaccine, etc.)	
🕅 Irinotecan		Genetic Counsel	ing	

Figure 59: Gall Bladder/Biliary Cancer Treatment Form

Field Name	Data Type	Description
)D	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
TxResect	Yes/No	Treatment - Resection
TxLap	Yes/No	Treatment - Laparoscopy
TxRadia	Yes/No	Treatment - Radiation
TXRadiaAdju	Number	Treatment - Radiation - Adjuvancy
TxChemo	Yes/No	Treatment - Chemo
TXChemoAdju	Number	Treatment - Chemo - Adjuvancy
TxChemoAVA	Yes/No	Treatment - Chemo - Avastin
TxChemoCap	Yes/No	Treatment - Chemo - Capecitabine
TxChemoErb	Yes/No	Treatment - Chemo - Erbitux
TxChemoFlu	Yes/No	Treatment - Chemo - Fluorouracil (5-FU)
TxChemoFUDR	Yes/No	Treatment - Chemo - FUDR
TxChemoGem	Yes/No	Treatment - Chemo - Gemcitabine
TxChemoIri	Yes/No	Treatment - Chemo - Irinotecan
TxChemoLeu	Yes/No	Treatment - Chemo - Leukovorin
TxChemoLev	Yes/No	Treatment - Chemo - Levamasole
TxChemoMit	Yes/No	Treatment - Chemo - Mitomycin
TxChemoOxa	Yes/No	Treatment - Chemo - Oxaliplatin
TxChemoTax	Yes/No	Treatment - Chemo - Taxol
TxChemoOth	Yes/No	Treatment - Chemo - Other
TxChemoOS	Text	Treatment - Chemo - Other - Specify
TxPal	Yes/No	Treatment - Palliation
TxPalRes	Yes/No	Treatment - Palliation - Pall, Resection
TxPalBypass	Yes/No	Treatment - Palliation - Bypass
TxPalCeliac	Yes/No	Treatment - Palliation - Celiac Block
TxPalPara	Yes/No	Treatment - Palliation - Paracentesis
TxPalTho	Yes/No	Treatment - Palliation - Thoracentesis
TxPalRad	Yes/No	Treatment - Palliation - Pall. Radiation
TxPalTrans	Yes/No	Treatment - Palliation - Transfusion
TxPalStens	Yes/No	Treatment - Palliation - Pall, Stenting
TxPalPV	Yes/No	Treatment - Palliation - PV Shunts
TxPalHAL	Yes/No	Treatment - Palliation - HAL
TxPalGasTube	Yes/No	Treatment - Palliation - Gastrostomy Tube
TxPalJejTube	Yes/No	Treatment - Palliation - Jejunstomy Tube
TxPalOth	Yes/No	Treatment - Palliation - Other
TxPalOS	Text	Treatment - Palliation - Other - Specify
TxExp	Yes/No	Treatment - Experimental protocol (ie. vaccine)
TxGene	Yes/No	Treatment - Gene Counseling

Figure 60: Gall Bladder/Biliary Cancer Treatment Table Schema

Microsoft Access - [GB_6a_Res : Form]
🗐 Eile Edit View Insert Format Records Tools Window Help Adobe PDF Type a question for help 🖌 🗗 🗙
⊻ - □ 1 ● □ ♥
- MS Sans Serif - 8 - B I U 画画画 公 - ム - ビー・
If Resection is Performed
Surgery Date of Admission: Date of Surgery: Procedure Type
🕅 Venous Resection 🕅 Venous Reconstruction 🕅 Arterial Resection 🕅 Arterial Reconstruction
Other Organs Resected: Estimated Blood Loss (cc): Transfusion If Yes, Units: Methods: FFP Cell Saver
Resection Attempt: © Successful © Unsuccessful - Reason: Post-Op Days in ICU: Post-Op Care Path: © Congruent © Divergent
🕅 NG/Gastrostomy Drainage > 7days 🛛 🕅 Abdominal Collection
🕅 Pulmonary Complications 🕅 Wound Infection 🔲 Leak
🗑 Liver Insufficiency (Total Bilirubin > 5) If Yes, Total Bilirubin:
Date of Discharge: Discharge Status:
Form View

Figure 61: Gall Bladder/Biliary Cancer Resection Form

Field Name	Data Type	Description	
ĴD	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
ResDAdm	Date/Time	Resection - Date of Admission	
ResDSurg	Date/Time	Resection - Date of Surgery	
ResPxType	Number	Resection - Procedure Type (Whipple, total pancreatectomy, distal pancreatectomy, etc)	
ResORTime	Number	Resection - OR Time (hr.)	
ResVenRes	Yes/No	Resection - Venous Resection	
ResVenRec	Yes/No	Resection - Venous Reconstruction	
ResArtRes	Yes/No	Resection - Arterial Resection	
ResArtRec	Yes/No	Resection - Arterial Reconstruction	
ResOrgans	Text	Resection - Other Organs Resection	
ResBloodLoss	Number	Resection - Estimated Blood Loss (cc)	
ResTransfusion	Yes/No	Resection - Tranfusion	
ResTUnits	Number	Resection - Transfusion Units	
ResTFFP	Yes/No	Resection - Transfusion - FFP	
ResTCell	Yes/No	Resection - Transfusion - Cell	
ResAttempt	Number	Resection - Resection Attempt	
ResAttemptUn	Number	Resection - Resection Unsuccessful Reason (Tumor involvement, Operative mishap, etc)	
ResPOCourse	Number	Resection - PO - Post-Op Care Path	
ResPODays	Number	Resection - PO - Time in ICU (days)	
ResPOInfection	Yes/No	Resection - PO - Wound infection	
ResPOLeak	Yes/No	Resection - PO - Leak	
ResPONG	Yes/No	Resection - PO - NG/gastrotomy drainage	
ResPOAbdominal	Yes/No	Resection - PO - Abdominal Collection	
ResPOPulmComp	Yes/No	Resection - PO - Pulminary Complications	
ResPOLiverInsuf	Yes/No	Resection - PO - Liver Insufficiency	
ResPOLiverTB	Number	Resection - PO - Liver Insufficiency - Total Bilirubin	
ResPODDischarge	Date/Time	Resection - Date of Discharge	
ResPODischStatus	Number	Resection - Discharge Status	

Figure 62: Gall Bladder/Biliary Cancer Resection Table Schema



Figure 63: Gall Bladder/Biliary Cancer No Resection Form

Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
NoResEvalDate	Date/Time	No Resection - Decision Date	
NoResNoHandle	Yes/No	No Resection - Couldn't Handle Proposed Treatment	
NoResRefused	Yes/No	No Resection - Refused Treatment	
NoResMagnitude	Yes/No	No Resection - Magnitude Not Worth Benefits	
NoResCeliacInvolve	Yes/No	No Resection - Celiac Trunk Involvement	
NoResSMAInvolve	Yes/No	No Resection - SMA Involvement	
NoResHepaticInvolve	Yes/No	No Resection - Hepatic Involvement	
NoResIVCInvolve	Yes/No	No Resection - Inferior Vena Cava Involvement	
NoResSMVInvolve	Yes/No	No Resection - SMV Involvement	
NoResPVInvolve	Yes/No	No Resection - Portal Vein Involvement	
NoResCirrhosis	Yes/No	No Resection - Cirrhosis	
NoResMetastatic	Yes/No	No Resection - Metastatic	

Figure 64: Gall Bladder/Biliary Cancer No Resection Table Schema

Microsoft Access - [GB_7_Path : Form]	
🔄 File Edit View Insert Format Records	ls Tools Window Help Adobe PDF Type a question for help → _ & ×
₩· ■ • ● Q ♥ X • • • • •	○ 😤 🛃 🏹 🖥 ▽ 🛤 ▷* 🕊 🖆 🕯 🕶 - 🔍 -
MS Sans Serif 🗾 10	• B I U ≣ ≣ ≣ 🏖 • 💁 • 🖉 • 🔲 • 📼 • .
Final Tumor Histology	
-	ח
(from best of imaging, FNA, pathology, etc)	
Pathology (if available)	
Tumor Size (cm): by	
TNM Staging: T N T	M R:
Form View	

Figure 65: Gall Bladder/Biliary Cancer Pathology Form

Field Name	Data Type	Description	
(D	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
Histology	Number	Histology	
ResPathT	Number	Resection - Pathology Staging - T	
ResPathN	Number	Resection - Pathology Staging - N	
ResPathM	Number	Resection - Pathology Staging - M	
ResPathR	Number	Resection - Pathology Staging - R	
ResPathV	Number	Resection - Pathology Staging - R	
ResPathSizeX	Number	Resection - Pathology Tumor Size (cm) - Width	
ResPathSizeY	Number	Resection - Pathology Tumor Size (cm) - Height	

Figure 66: Gall Bladder/Biliary Cancer Pathology Table Schema
🖉 Microsoft Access - [Follow-up I	nformation]	- 🗆 🗵
📅 File Edit View Insert Fo	rmat <u>Records Tools Window H</u> elp Adobe PDF Type a question for help	8 ×
🔟 • 🔲 🖬 🖨 🖪 🖤 🐰	ⓑ í ⋈ 🍓 🛃 🏹 🏹 🗑 🗸 🛤 🕨 🗰 🗇 🔍 .	
Gall Blauder/Biliary		
Follow-up Informatio	n Add Record Find Record Delete Record	
MR #Name?	Follow-up Window:	
Date of Visit:	Weight (pounds):	
ECOG performance status: © 0	@ 1 @ 2 @ 3 @ 4	
Lab Value:		
	Alkaline Phosphotase	
CA19-9: Total B		
Redeveloped Symptoms:		
🕅 Weight loss	🏾 Biliary colic 🛛 🕅 Pruritis 🕅 Back pain	
how much (pounds)	🕅 Nausea 🛛 🕅 Abdominal pain 🕅 Indigestion	
🕅 Jaundice	🕅 Vomiting 🕅 Other Specify:	
🗰 Cholecystitis	💹 Clay colored stool	
🕅 Cholangitis	🕅 Fatigue	
Status:		
C Died Death Date:		
O N.E.D.		
C A.W.D., Method of Detection:	🏾 Lab 🛛 🕅 Radiologic Evidence 🕅 Clinical Evidence	
1		
Form View		1

Figure 67: Gall Bladder/Biliary Cancer Follow-Up Form

	Field Name	Data Type	Description
ID	00000000000000000000000000000000000000	AutoNumber	· · · · · · · · · · · · · · · · · · ·
MR		Text	Meditech Medical Record Number for Patient
FUV	Vin	Number	Follow-Up Windows
Visit	:Date	Date/Time	Visit Date
Wei	ight	Number	Weight (lbs.)
QOL	Lscore	Number	QoL Score (0-100)
ECC	DG	Number	ECOG Score (0-4)
Labo	CEA	Number	Laboratory - CEA
Lab	CA19-9	Number	Laboratory - CA19-9
Lab	Alb	Number	Laboratory - Albumin
Lab	Bili	Number	Laboratory - Bilirubin
Lab	Alka	Number	Laboratory - Alkaline phosphotase
SxW	Vtloss	Yes/No	Symptoms - Weight Loss
SxW	VtlossP	Number	Symptoms - Weight Loss (lbs.)
SxJa	aun	Yes/No	Symptoms - Jaundice
SxC	Thole	Yes/No	Symptoms - Cholecystitis
SxC	:hola	Yes/No	Symptoms - Cholangitis
SxB	IC	Yes/No	Symptoms - Biliary Colic
SxN	lau	Yes/No	Symptoms - Nausea
SxV	'om	Yes/No	Symptoms - Vomiting
SxC	CS	Yes/No	Symptoms - Clay Colored Stool
SxF.	ati	Yes/No	Symptoms - Fatigue
SxP	'ru	Yes/No	Symptoms - Pruritis
SxIr	nd	Yes/No	Symptoms - Indigestion
SxA	vbd	Yes/No	Symptoms - Abdominal Pain
SxB.	lack	Yes/No	Symptoms - Back Pain
15x0)T	Yes/No	Symptoms - Other
SxO	TSpe	Text	Symptoms - Other - Specify
Stat		Number	Status (NED, AWD, Died)
Dea	athDate	Date/Time	Death Date
Stat	tusAWDLab	Yes/No	AWD - Lab Evidence
Stat	tusAWDRad	Yes/No	AWD - Radiology Evidence
Stat	tusAWDCli	Yes/No	AWD - Clinical Evidence

Figure 68: Gall Bladder/Biliary Cancer Follow-Up Table Schema

3.1.4 Gastric Cancer

¥• 🖬 🖬 🖨 🖪 💖 🐰 🎙	A 10 0 4	I VAN	🔥)+	- · ?	
MS Sans Serif			2 · A · 2 ·	Participant in the second s	
resumptive Diagnosis				•	
esumplive Diagnosis		Care			
	·				
<u>resentation</u>					
Date of Evaluation ECOG Perfo	rmance Status	Height (in.) Wei	ht (lbs.)	1	
	@ 2 @ 3 @ 4				
Sumptome					
	🐺 Biliary Colic	m Pruritis	Indigestion		
Weight Loss	I₩ Biliary Colic I₩ Nausea	🕅 Pruritis 🗑 Abdominal Pain	Indigestion		
Weight Loss How Much (pounds):	🕅 Nausea	🗑 Abdominal Pain	🕅 Dysphagia		
Weight Loss How Much (pounds):	I₩ Nausea I₩ Vomiting	🕅 Abdominal Pain 🕅 Back Pain			
Symptoms Weight Loss How Much (pounds): Jaundice Cholecystitis Cholangitis	🕅 Nausea	🕅 Abdominal Pain 🕅 Back Pain	🕅 Dysphagia		

Figure 69: Gastric Cancer Presentation Form

Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)	
DemEvalDate	Date/Time	Demographics - Date Evaluated by Surgical Oncology	
DemECOG	Number	Demographics - ECOG Score (0-4)	
DemHeight	Number	Demographics - Height in Inches of Patient	
DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission	
SxWtloss	Yes/No	Initial Symptoms - Weight Loss	
SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds	
SxJaun	Yes/No	Initial Symptoms - Juandice	
SxChole	Yes/No	Initial Symptoms - Cholecystitis	
SxChola	Yes/No	Initial Symptoms - Cholangitis	
SxBC	Yes/No	Initial Symptoms - Biliary Colic	
SxNau	Yes/No	Initial Symptoms - Nausea	
SxVom	Yes/No	Initial Symptoms - Vomiting	
SXCCS	Yes/No	Initial Symptoms - Clay Colored Stool	
SxFati	Yes/No	Initial Symptoms - Fatigue	
SxPru	Yes/No	Initial Symptoms - Pruritis	
SxInd	Yes/No	Initial Symptoms - Indigestion	
SxAbd	Yes/No	Initial Symptoms - Abdominal Pain	
SxBack	Yes/No	Initial Symptoms - Back Pain	
SxDyspha	Yes/No	Initial Symptoms - Dysphagia	
SxSatiety	Yes/No	Initial Symptoms - Early Satiety	
SXOT	Yes/No	Initial Symptoms - Other	
SxOTSpe	Text	Initial Symptoms - Other - Specify	

Figure 70: Gastric Cancer Presentation Table Schema

▶ • • • • • • • • • • • • • • • • • • •			
Respiratory Renal Failure G Le Hypertension Renal Disorder Directory	Failure/Cirrhosis Failure / Cirrhosis Failure / Cirrhosis Failure / Cirrhosis sthan Six Months Father Dx: eater than Six Months Mother Dx: eater than Six Months et Control Relation: Relation: Relation control R	Radiation F Surgery	

Figure 71: Gastric Cancer Medical History Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
CxHF	Yes/No	Comorbidities - Heart Failure
CxIHD	Yes/No	Comorbidities - Ischemic Heart Disease
CxResp	Yes/No	Comorbidities - Respiratory
CxDiab	Yes/No	Comorbidities - Diabetes
CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral
CxDiabDiet	Yes/No	Comorbidities - Diabetes - Insulin - Diet Control
CxDiabOnset	Number	Comorbidities - Diabetes - Onset (1=Less than six months, 2 =Greater than six months)
CXRF	Yes/No	Comorbidities - Renal Failure
CxHyper	Yes/No	Comorbidities - Hypertension
CxBleed	Yes/No	Comorbidities - Bleeding Disorder
CxLiver	Yes/No	Comorbidities - Liver Failure
CxMal	Yes/No	Comorbidities - Malnutrition
CxPriorCancer	Number	Comorbidities - Prior Cancer Dx
CxPriorCancerChemo	Yes/No	Comorbidities - Prior Cancer Dx - Chemo
CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation
CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery
SHCigarette	Yes/No	Social History - Cigarettes (significant use)
SHAlcohol	Yes/No	Social History - Alcohol (significant use)
SHDrugUse	Yes/No	Social History - Drug Use
SHExposure	Yes/No	Social History - Environmental Exposure
SHOther	Yes/No	Social History - Other
SHOtherS	Text	Social History - Other - Specify
FamilyFatherDx	Number	Family History - Father Dx
FamilyMotherDx	Number	Family History - Mother Dx
FamilyOther1	Text	Family History - Other1
FamilyOther1Dx	Number	Family History - Other1 Dx
FamilyOther2	Text	Family History - Other2
FamilyOther2Dx	Number	Family History - Other2 Dx

Figure 72: Gastric Cancer Medical History Table Schema

Microsoft Access -	[Gas_3a_Serum	: Form]					_ 🗆 ×
🗐 File Edit View	Insert Format	Records I	ools <u>W</u> indow	v <u>H</u> elp Ado <u>b</u> e F	PDF	Type a question for help	×
🔟 • 🖬 🔁 🍯	Q.♥ % B	B 10 8		y 7 7 M)* W 😭	🗗 🔄 🖉 💶	
Contraction of the second seco	S Sans Serif	1140 22 1 1 1 1 1 1 1 1 1	BI	u F = =	2 · A · .	<u>/</u> · <u>·</u> · <u>·</u> ·	
Serum Studies							
CEA:	Albumin		ALK	ALT			
CA19-9:	Total Bilirubin		AST:	Amyla	se:		
H. Pylori Status: C Po:	sitive C Negative						
Form View							

Figure 73: Gastric Cancer Serum Studies Form

1	Field Name	Data Type	Description	
81	D	AutoNumber	ID	2
	MR	Text	Meditech Medical Record Number for Patient	
1	LabCEA	Number	Laboratory - CEA	
	LabCA19-9	Number	Laboratory - CA19-9	
8	LabAlb	Number	Laboratory - Albumin	
	LabBili	Number	Laboratory - Bilirubin	
1	LabAlka	Number	Laboratory - Alkaline phosphotase	
	LabALT	Number	Laboratory - ALT	
	LabAST	Number	Laboratory - AST	
	LabAmylase	Number	Laboratory - Amylase	
1	LabHPylori	Number	Laboratory - H. Pylori Status	

Figure 74: Gastric Cancer Serum Studies Table Schema

¥ · ■ • ● G ♥ ↓ • •		
Piagnostic Imaging Procedur		
CT with Panreatic Protocol/CT/		
Date of Procedure:	2.	
🐷 Celiac Artery Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
🗮 Superior Mesenteric Artery Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
👿 Hepatic Artery Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
👿 Inferior Vena Cava Involvement	C Open C Abutted C Encased C Occluded C Unknown	
🕅 Superior Mesenteric Vein Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
Rortal Vein Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown	
Nodes: 🕅 Celiac Nodal Disease	👿 Other Nodal Disease 🛛 🕅 No Nodal Assessment or Mention	
Tumor Size (cm): by		
Chest X-Ray (CXR)		
Percutaneous Transhepatic Ch	polangiography (PTC)	
Date of Procedure:		
Stenting Type: @ Internal	@ External	
· · · · · · · · · · · · · · · · · · ·		

Figure 75: Gastric Cancer Diagnostic Imaging Form

	Field Name	Data Type	Description
3)	ID	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
	CXRDx	Yes/No	CXR - Diagnosis
	CTDx	Yes/No	CT - Diagnosis
	CTEvalDate	Date/Time	CT - Date Evaluated
	CTVascOmit	Yes/No	CT - Vascular Omission
	CTCeliac	Yes/No	CT - Celiac Involvement
	CTCeliacClass	Number	CT - Celiac Involvement Class
	CTSMA	Yes/No	CT - SMA Involvement
	CTSMAClass	Number	CT - SMA Involvement Class
	CTHepatic	Yes/No	CT - Hepatic Involvement
	CTHepaticClass	Number	CT - Hepatic Involvement
	CTInferior	Yes/No	CT - Inferior Vena Cava Involvement
	CTInferiorClass	Number	CT - Inferior Vena Cava Involvement Class
	CTSMV	Yes/No	CT - SMV Involvement
	CTSMVClass	Number	CT - SMV Involvement Class
	CTPortal	Yes/No	CT - Portal Vein Involvement
	CTPortalClass	Number	CT - Portal Vein Involvement Class
	CTCeliacNode	Yes/No	CT - Celiac Nodal Disease
	CTOtherNode	Yes/No	CT - Other Nodal Disease
	CTNodeOmit	Yes/No	CT - Node Omission
	CTTumorSizeX	Number	CT - Tumor Size (cm) - Width
	CTTumorSizeY	Number	CT - Tumor Size (cm) - Height
	PTCDx	Yes/No	PTC - Diagnosis
	PTCEvalDate	Date/Time	PTC - Date Evaluated
	PTCStent	Yes/No	PTC - Stent
	PTCStentType	Number	PTC - Stent Type

Figure 76: Gastric Cancer Diagnostic Imaging Table Schema



Figure 77: Gastric Cancer Preliminary Outlook Form

S.	Field Name	Data Type	Description
81	ID	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
1	PreOutlook	Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic or Equivocal Findi

Figure 78: Gastric Cancer Preliminary Outlook Table Schema

¥ • 🖬 •3 (4 5 (2 °) × •6 €.	∽
reatment Course	
『 <u>Resection</u> 『 <u>Radiation</u> 『 <u>Adjuvant</u> © Neoadjuvant © Both 『 <u>Chemotherapy</u>	Staging Laparoscopy/Laparotomy Palliative Measures Bypass HAL Gastrostomy Tube PV Shunts
Adjuvant Neoadjuvant Both Avastin Capecitabine Erbitux Mitomycin Fluorouracil (5-FU) Capel Taxol	I Jejunstomy Tube I Pall. Stenting III Celiac Block. III Pall. Radiation III Paracentesis III Pall. Resection III Thoracentesis III Other - Specify. IIII Transfusion
Image: Gemeitabine Image: Other - Specify: Image: Innotecan Image: Other - Specify:	Experimental Protocol (vaccine, etc.) Genetic Counseling

Figure 79: Gastric Cancer Treatment Form

	Field Name	Data Type	Description				
2.	(D	AutoNumber	ID				
1	MR	Text	Meditech Medical Record Number for Patient				
Î	TxResect Yes/No		Treatment - Resection				
-	TxLap Yes/No		Treatment - Laparoscopy				
-	TxRadia Yes/No		Treatment - Radiation				
	TXRadiaAdiu Number		Treatment - Radiation - Adjuvancy				
1	TxChemo Yes/No TXChemoAdiu Number		Treatment - Chemo				
-			Treatment - Chemo - Adjuvancy				
-	TxChemoAVA	Yes/No	Treatment - Chemo - Avastin				
	TxChemoCap	Yes/No	Treatment - Chemo - Capecitabine				
1	TxChemoErb	Yes/No	Treatment - Chemo - Erbitux				
-	TxChemoFlu	Yes/No	Treatment - Chemo - Fluorouracil (5-FU)				
-	TxChemoFUDR	Yes/No	Treatment - Chemo - FUDR				
1	TxChemoGem	Yes/No	Treatment - Chemo - Gemcitabine				
1	TxChemoIri	Yes/No	Treatment - Chemo - Trinotecan				
3	TxChemoLeu	Yes/No	Treatment - Chemo - Leukovorin				
-	TxChemoLev	Yes/No	Treatment - Chemo - Levamasole				
7	TxChemoMit Yes/No TxChemoOxa Yes/No TxChemoTax Yes/No		Treatment - Chemo - Mitomycin				
1			Treatment - Chemo - Oxaliplatin				
3			Treatment - Chemo - Taxol				
-	TxChemoOth	Yes/No	Treatment - Chemo - Other				
	TxChemoOS	Text	Treatment - Chemo - Other - Specify				
-	TxPal	Yes/No	Treatment - Palliation				
-	TxPalRes	Yes/No	Treatment - Palliation - Pall, Resection				
-	TxPalBypass	Yes/No	Treatment - Pallation - Bypass				
-		Yes/No	Treatment - Pallation - Celiac Block				
-	TxPalPara	Yes/No	Treatment - Pallistion - Paracentesis				
- 1	TxPalTho	Yes/No	Treatment - Pallation - Thoracentesis				
-	TxPalRad	Yes/No	Treatment - Pallation - Pall. Radiation				
-		Yes/No	Treatment - Pallation - Transfusion				
-	TxPalStens	Yes/No	Treatment - Palliation - Pall, Stenting				
	TxPalPV	Yes/No	Treatment - Pallation - PV Shurts				
1000	TxPalHAL	Yes/No	Treatment - Palliation - HAL				
-	TXPalGasTube	Yes/No	Treatment - Palliation - Gastrostomy Tube				
100	TxPalJeiTube	Yes/No	Treatment - Palliation - Jejunstomy Tube				
	TxPalOth	Yes/No	Treatment - Palliation - Jejunstomy Tube				
100	TxPalOS	Text	Treatment - Palliation - Other - Specify				
	TXEXD	Yes/No					
-	TxExp		Treatment - Experimental protocol (ie. vaccine)				
4	IXGene	Yes/No	Treatment - Gene Counseling				

Figure 80: Gastric Cancer Treatment Table Schema

Microsoft Acc Zoom Gas_6a_Res : Form]	-OX
Eg Eile Edit View Insert Format Records Iools Window Help Adobe PDF Type a question for help	×
<u>₩</u> • <u>₩</u>	
M5 Sans Serif • 8 • B I U ≣ ≣ ≣ 2 • ▲ • 2 • • • • •	
If Resection is Performed	-
Surgery Date of Admission: Date of Surgery: Procedure Type Image: Comparison of Com	
🗑 Venous Resection 🖉 Venous Reconstruction 🖉 Arterial Resection 🖉 Arterial Reconstruction	
Other Organs Resected: Estimated Blood Loss (cc):	
🕅 Transfusion If Yes, Units: 🛛 🛛 Methods: 🕅 FFP 🕅 Cell Saver	
Resection Attempt: If Successful Image: Successful - Reason: Image: Successful - Reason: Post-Op Image: Successful - Reason: Days in ICU: Image: Successful - Reason: Post-Op Care Path: Image: Congruent Divergent Image: Successful - Reason:	
🕅 NG/Gastrostomy Drainage > 7days 🔲 Abdominal Collection 🖉 Prolonged Ileus	
🖉 Pulmonary Complications 🖉 Wound Infection 🖉 Leak 🖉 Small Bowel Obstruction	
🕼 Liver Insufficiency (Total Bilirubin > 5) If Yes, Total Bilirubin:	
Date of Discharge:	
	1
Form View	

Figure 81: Gastric Cancer Resection Form

Field Name	Data Type	Description			
[D	AutoNumber	ID			
MR	Text	Meditech Medical Record Number for Patient			
ResDAdm Date/Time		Resection - Date of Admission			
ResDSurg Date/Time		Resection - Date of Surgery			
ResPxType Number		Resection - Procedure Type (Whipple, total pancreatectomy, distal pancreatectomy, etc)			
ResORTime	Number	Resection - OR Time (hr.)			
ResVenRes	Yes/No	Resection - Venous Resection			
ResVenRec	Yes/No	Resection - Venous Reconstruction			
ResArtRes	Yes/No	Resection - Arterial Resection			
ResArtRec	Yes/No	Resection - Arterial Reconstruction			
ResOrgans	Text	Resection - Other Organs Resection			
ResBloodLoss	Number	Resection - Estimated Blood Loss (cc)			
ResTransfusion	Yes/No	Resection - Tranfusion			
ResTUnits	Number	Resection - Transfusion Units			
ResTFFP	Yes/No	Resection - Transfusion - FFP			
ResTCell	Yes/No	Resection - Transfusion - Cell			
ResAttempt	Number	Resection - Resection Attempt Resection - Resection Unsuccessful Reason (Tumor involvement, Operative mishap, etc)			
ResAttemptUn	Number				
ResPOCourse	Number	Resection - PO - Post-Op Care Path			
ResPODays	Number	Resection - PO - Time in ICU (days)			
ResPOInfection	Yes/No	Resection - PO - Wound infection			
ResPOLeak	Yes/No	Resection - PO - Leak			
ResPONG	Yes/No	Resection - PO - NG/gastrotomy drainage			
ResPOAbdominal	Yes/No	Resection - PO - Abdominal Collection			
ResPOPulmComp	Yes/No	Resection - PO - Pulminary Complications			
ResPOLiverInsuf	Yes/No	Resection - PO - Liver Insufficiency			
ResPOLiverTB	Number	Resection - PO - Liver Insufficiency - Total Bilirubin			
ResPODDischarge	Date/Time	Resection - Date of Discharge			
ResPODischStatus	Number	Resection - Discharge Status			

Figure 82: Gastric Cancer Resection Table Schema

	_6b_NoRes : Form]				×
🗃 Eile Edit View Ins	ert F <u>o</u> rmat <u>R</u> ecords	s <u>T</u> ools <u>W</u> indow <u>H</u> el	p Ado <u>b</u> e PDF	Type a question for help	8 >
¥ · 🖬 🗞 🖨 🖪 '	🎔 👗 🖻 💼 🗠		T M >* M	8 6 10 .	
*	*	1.0000	F = = <u>></u> - <u>A</u>	- 2	
Resection is N	ot Performed	1			
Date of Decision:					
Reasons (select all th	nat apply):				
Clinical Decision					
🕅 Patient Couldn't Hand	le Proposed Treatment				
Patient Refused Treat	ment				
Proposed Magnitude Risks Not Worth Like	of Treatment and y Benefit				
Vascular Involvement		Additional Disease	2		
	ient	🕅 Cirrhosis			
🔲 🕅 Celiac Artery Involven			ana		
 Celiac Artery Involven Superior Mesenteric A 	rtery Involvement	🔟 Evidence of Me	stastasis		
		I≋ Evidence of Me	etastasis		
Superior Mesenteric A	ement	I≋ Evidence of Me	etastasis		
Superior Mesenteric A	ement volvement	I₩ Evidence of Me	erastasis		

Figure 83: Gastric Cancer No Resection Form

Field Name	Data Type	Description		
ID	AutoNumber	ID		
MR	Text	Meditech Medical Record Number for Patient		
NoResEvalDate	Date/Time	No Resection - Decision Date		
NoResNoHandle	Yes/No	No Resection - Couldn't Handle Proposed Treatment		
NoResRefused	Yes/No	No Resection - Refused Treatment		
NoResMagnitude	Yes/No	No Resection - Magnitude Not Worth Benefits		
NoResCeliacInvolve	Yes/No	No Resection - Celiac Trunk Involvement		
NoResSMAInvolve	Yes/No	No Resection - SMA Involvement		
NoResHepaticInvolve	Yes/No	No Resection - Hepatic Involvement		
NoResIVCInvolve	Yes/No	No Resection - Inferior Vena Cava Involvement		
NoResSMVInvolve	Yes/No	No Resection - SMV Involvement		
NoResPVInvolve	Yes/No	No Resection - Portal Vein Involvement		
NoResCirrhosis	Yes/No	No Resection - Cirrhosis		
NoResMetastatic	Yes/No	No Resection - Metastatic		

Figure 84: Gastric Cancer No Resection Table Schema

Microsoft Access - [Ga	s_7_Path : Form]						- 🗆 ×
NAME AND ADDRESS OF AD	nsert F <u>o</u> rmat <u>R</u> e	ecords <u>T</u> ools	s <u>W</u> indow	Help Adobe F	PDF	Type a question for help	×
	💖 🐰 🖻 💼	10		* 🔁 🗸 🕅)* W 😭	d 'a • Q .	
						<u>_</u> . <u> </u> . <u> </u> .	
Final Tumor Hi	stology						
1		–					
(from best of imaging, FNA,	pathology, etc)						
	1-11 1-11-11-11-11-11-11-11-11-11-11-11-						
Pathology (if av	/ailable)						
Tumor Size (cm):	by						
TNM Staging: T	- N	J M [_	R:	1		
Form View							

Figure 85: Gastric Cancer Pathology Form

Field Name	Data Type	Description			
ID MR	AutoNumber	ID			
MR	Text	Meditech Medical Record Number for Patient			
Histology	Number	Histology			
ResPathT	Number	Resection - Pathology Staging - T			
ResPathN	Number	Resection - Pathology Staging - N			
ResPathM	Number	Resection - Pathology Staging - M			
ResPathR	Number	Resection - Pathology Staging - R			
ResPathV	Number	Resection - Pathology Staging - R			
ResPathSizeX	Number	Resection - Pathology Tumor Size (cm) - Width			
ResPathSizeY	Number	Resection - Pathology Tumor Size (cm) - Height			

Figure 86: Gastric Cancer Pathology Table Schema

Microsoft Access - [Follow-up Information]	X
Ele Edit Yiew Insert Format Records Tools Window Help Adobe PDF Type a question for help -	×
Gastric Tul Print (hp deskjet 920c series)	
Follow-up Information Add Record Find Record Delete Record	
MR Name? Follow-up Window:	
Date of Visit: QOL score: QOL score:	
ECOG performance status: © 0 © 1 © 2 © 3 © 4	
Lab Value:	
CEA: Albumin Alkaline Phosphotase	
CA19-9: Total Bilirubin	
Redeveloped Symptoms:	
🏾 Weight loss 🖉 Biliary colic 🖉 Prunitis 🖉 Back pain	
how much (pounds) 🛛 🕅 Nausea 🕅 Abdominal pain 🕅 Indigestion	
I Jaundice Vomiting Other Specify:	
Im Cholecystitis Im Clay colored stool	
The Cholangitis The Fatigue	
Status:	
C Died Death Date:	
O NED.	
C A.W.D., Method of Detection: 🔲 Lab 📁 Radiologic Evidence 🛒 Clinical Evidence	
Form View	11

Figure 87: Gastric Cancer Follow-Up Form

12	Field Name	Data Type	Description			
I		AutoNumber	76 N N D7 00 12 76 70			
M	1R	Text	Meditech Medical Record Number for Patient			
F	UWin	Number	Follow-Up Windows			
V	'isitDate	Date/Time	Visit Date			
W	Veight	Number	Weight (lbs.)			
Ç	OLscore	Number	QoL Score (0-100)			
E	COG	Number	ECOG Score (0-4)			
L	abCEA	Number	Laboratory - CEA			
L	abCA19-9	Number	Laboratory - CA19-9			
L	abAlb	Number	Laboratory - Albumin			
L	abBili	Number	Laboratory - Bilirubin			
L	abAlka	Number	Laboratory - Alkaline phosphotase			
5	ixWtloss	Yes/No	Symptoms - Weight Loss			
S	ixWtlossP	Number	Symptoms - Weight Loss (lbs.)			
S	ixJaun	Yes/No	Symptoms - Jaundice			
S	ixChole	Yes/No	Symptoms - Cholecystitis			
s	ixChola	Yes/No	Symptoms - Cholangitis			
S	бхВС	Yes/No	Symptoms - Biliary Colic			
5	ixNau	Yes/No	Symptoms - Nausea			
S	ixVom	Yes/No	Symptoms - Vomiting			
S	ixCC5	Yes/No	Symptoms - Clay Colored Stool			
S	ixFati	Yes/No	Symptoms - Fatigue			
S	ixPru	Yes/No	Symptoms - Pruritis			
S	ixInd	Yes/No	Symptoms - Indigestion			
s	ixAbd	Yes/No	Symptoms - Abdominal Pain			
S	ixBack	Yes/No	Symptoms - Back Pain			
S	ixOT	Yes/No	Symptoms - Other			
S	ixOTSpe	Text	Symptoms - Other - Specify			
S	itatus	Number	Status (NED, AWD, Died)			
D	eathDate	Date/Time	Death Date			
S	itatusAWDLab	Yes/No	AWD - Lab Evidence			
Is	itatusAWDRad	Yes/No	AWD - Radiology Evidence			
S	itatusAWDCli	Yes/No	AWD - Clinical Evidence			

Figure 88: Gastric Cancer Follow-Up Table Schema

3.1.5 Esophageal Cancer

		Window Help Adobe	I THE REPORT OF A DESCRIPTION OF A DESCRIPANTO OF A DESCRIPTION OF A DESCRIPTION OF A DESCRIPTION OF A DESCR	ion for help 💌 🗕 🗗
· 🖬 🔁 🖨 🖪 💖 🐰	₽ CB 10 GB 24	👬 🏹 🔁 🗸 🕅) >* × 😰 🗗 🖞	- Q.
MS Sans Serif	• 8 • B	/ ⊻ ≡ ≡ ≡	<u></u> → <u></u> + <u></u> +	-
esumptive Diagnos	is at Onset of	Care		
ovaniparo blagnov		<u></u>		
	<u> </u>			
esentation				
esemanon				
		Height (in.) Weigh	<u>nt (lbs.)</u>	
@ 0 @ ·	1 @ 2 @ 3 @ 4			
<u>ymptoms</u>				
🖩 Weight Loss	🕅 Biliary Colic	🕅 Pruritis	Indigestion	
How Much (pounds):	🕅 Nausea	🕅 Abdominal Pain	🕅 Dysphagia	
🖩 Jaundice	🕅 Vomiting	🕅 Back Pain	🕅 Early Satiety	
	🕅 Clay Colored Stool	🕅 Other Specify:		
Cholecystitis				
 Cholecystitis Cholangitis 	🕅 Fatigue			

Figure 89: Esophageal Cancer Presentation Form

Field Name	Data Type	Description					
ID	AutoNumber	ID					
MR	Text	Meditech Medical Record Number for Patient					
PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)					
DemEvalDate	Date/Time	Demographics - Date Evaluated by Surgical Oncology					
DemECOG	Number	Demographics - ECOG Score (0-4)					
DemHeight	Number	Demographics - Height in Inches of Patient					
DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission					
SxWtloss	Yes/No	Initial Symptoms - Weight Loss					
SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds					
SxJaun	Yes/No	Initial Symptoms - Juandice					
SxChole	Yes/No	Initial Symptoms - Cholecystitis					
SxChola	Yes/No	Initial Symptoms - Cholangitis					
SxBC	Yes/No	Initial Symptoms - Biliary Colic					
SxNau	Yes/No	Initial Symptoms - Nausea					
SxVom	Yes/No	Initial Symptoms - Vomiting					
SxCCS	Yes/No	Initial Symptoms - Clay Colored Stool					
SxFati	Yes/No	Initial Symptoms - Fatigue					
SxPru	Yes/No	Initial Symptoms - Pruritis					
SxInd	Yes/No	Initial Symptoms - Indigestion					
SxAbd	Yes/No	Initial Symptoms - Abdominal Pain					
SxBack	Yes/No	Initial Symptoms - Back Pain					
SxDyspha	Yes/No	Initial Symptoms - Dysphagia					
SxSatiety	Yes/No	Initial Symptoms - Early Satiety					
SXOT	Yes/No	Initial Symptoms - Other					
SxOTSpe	Text	Initial Symptoms - Other - Specify					

Figure 90: Esophageal Cancer Presentation Table Schema

Microsoft Access - [Eso_2]	History : Form]		<u>_</u> O×
Image: Split Spli		Window Help Adobe PDF I	Type a question for help - <
	Malnutrition Liver Failure/Cirrhosis Diabetes C Less than Six Months Greater than Six Months Oral Agents Diet Control ar Drug Use rmental Exposure	Cancer History Patient Prior Dx: Image: Chemo Image: Radiation Father Dx: Image: Radiation Mother Dx: Image: Radiation Other Relation: Image: Radiation Related Dx: Image: Radiation Related Dx: Image: Radiation	✓ I Surgery ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓
Form View			

Figure 91: Esophageal Cancer Medical History Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
CxHF	Yes/No	Comorbidities - Heart Failure
CxIHD	Yes/No	Comorbidities - Ischemic Heart Disease
CxResp	Yes/No	Comorbidities - Respiratory
CxDiab	Yes/No	Comorbidities - Diabetes
CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral
CxDiabDiet	Yes/No	Comorbidities - Diabetes - Insulin - Diet Control
CxDiabOnset	Number	Comorbidities - Diabetes - Onset (1=Less than six months, 2 =Greater than six months)
CxRF	Yes/No	Comorbidities - Renal Failure
CxHyper	Yes/No	Comorbidities - Hypertension
CxBleed	Yes/No	Comorbidities - Bleeding Disorder
CxLiver	Yes/No	Comorbidities - Liver Failure
CxMal	Yes/No	Comorbidities - Malnutrition
CxPriorCancer	Number	Comorbidities - Prior Cancer Dx
CxPriorCancerChemo	Yes/No	Comorbidities - Prior Cancer Dx - Chemo
CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation
CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery
SHCigarette	Yes/No	Social History - Cigarettes (significant use)
SHAlcohol	Yes/No	Social History - Alcohol (significant use)
SHDrugÜse	Yes/No	Social History - Drug Use
SHExposure	Yes/No	Social History - Environmental Exposure
SHOther	Yes/No	Social History - Other
SHOtherS	Text	Social History - Other - Specify
FamilyFatherDx	Number	Family History - Father Dx
FamilyMotherDx	Number	Family History - Mother Dx
FamilyOther1	Text	Family History - Other1
FamilyOther1Dx	Number	Family History - Other1 Dx
FamilyOther2	Text	Family History - Other2
FamilyOther2Dx	Number	Family History - Other2 Dx

Figure 92: Esophageal Cancer Medical History Table Schema

PMicrosoft Acc	ess - [Eso_3a_9	Serum : Form]			
EB Eile Edit	View Insert	Format <u>R</u> ecord	ls <u>T</u> ools <u>W</u> indo	ow Help Adobe PDF	Type a question for help 🔽 🗕 🗗 🗙
🛛 🖌 🖌 🔛	a 🛛 💱	X 🖻 💼 片		V T V M >* N	< 🗗 🗗 🔄 🕄 🚬
E.	MS Sans Serif	• 8	• B I	U 📰 🗏 💆 🗸	▲ • ⊿ • □ • □ • .
Serum Stud	lies				
CEA:	Albumin		ALK	ALT	
CA19-9:	Total Bilir	rubin [AST:	Amylase:	
H. Pylori Status: 🤿	Positive 🤆 Ne	egative			
					71
Form View					1.

Figure 93: Esophageal Cancer Serum Studies Form

Field Name	Data Type	Description	
Ĩ▶ ID MR	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
LabCEA	Number	Laboratory - CEA	
LabCA19-9	Number	Laboratory - CA19-9	
LabAlb	Number	Laboratory - Albumin	
LabBili	Number	Laboratory - Bilirubin	
LabAlka	Number	Laboratory - Alkaline phosphotase	
LabALT	Number	Laboratory - ALT	
LabAST	Number	Laboratory - AST	
LabAmylase	Number	Laboratory - Amylase	
LabHPylori	Number	Laboratory - H. Pylori Status	

Figure 94: Esophageal Cancer Serum Studies Table Schema

Ele Edit View Insert Format I		Type a question for help
😫 • 日 🛍 🍯 🖪 🖤 👗 🖻 🖡	The second s	
*	→ B Z U E E E E A + A	<u>A</u> - <u>2</u> - <u>-</u>
iagnostic Imaging Procedur	es	
CT with Panreatic Protocol/CT/	N N	
Date of Procedure:	T	
Eliac Artery Involvement	C Open C Abutted C Encased C Occluded C	Unknown
🕅 Superior Mesenteric Artery Involvement	C Open C Abutted C Encased C Occluded C	Unknown
🕅 Hepatic Artery Involvement	C Open C Abutted C Encased C Occluded C	Unknown
🕅 Inferior Vena Cava Involvement	C Open C Abutted C Encased C Occluded C	Unknown
📓 Superior Mesenteric Vein Involvement	@ Open @ Abutted @ Encased @ Occluded @	Unknown
🕅 Portal Vein Involvement	C Open C Abutted C Encased C Occluded C	Unknown
Nodes: 🔲 Celiac Nodal Disease	🛛 🕅 Other Nodal Disease 🛛 🕅 No Nodal Assessment	or Mention
Tumor Size (cm): by		
Chest X-Ray (CXR)		
Percutaneous Transhepatic Ch	volengingraphy (PTC)	
Date of Procedure:		
Stenting Type: @ Internal	@ External	

Figure 95: Esophageal Cancer Diagnostic Imaging Form

Field Name	Data Type	Description
D ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
CXRDx	Yes/No	CXR - Diagnosis
CTDx	Yes/No	CT - Diagnosis
CTEvalDate	Date/Time	CT - Date Evaluated
CTVascOmit	Yes/No	CT - Vascular Omission
CTCeliac	Yes/No	CT - Celiac Involvement
CTCeliacClass	Number	CT - Celiac Involvement Class
CTSMA	Yes/No	CT - SMA Involvement
CTSMAClass	Number	CT - SMA Involvement Class
CTHepatic	Yes/No	CT - Hepatic Involvement
CTHepaticClass	Number	CT - Hepatic Involvement
CTInferior	Yes/No	CT - Inferior Vena Cava Involvement
CTInferiorClass	Number	CT - Inferior Vena Cava Involvement Class
CTSMV	Yes/No	CT - SMV Involvement
CTSMVClass	Number	CT - SMV Involvement Class
CTPortal	Yes/No	CT - Portal Vein Involvement
CTPortalClass	Number	CT - Portal Vein Involvement Class
CTCeliacNode	Yes/No	CT - Celiac Nodal Disease
CTOtherNode	Yes/No	CT - Other Nodal Disease
CTNodeOmit	Yes/No	CT - Node Omission
CTTumorSizeX	Number	CT - Tumor Size (cm) - Width
CTTumorSizeY	Number	CT - Tumor Size (cm) - Height
PTCDx	Yes/No	PTC - Diagnosis
PTCEvalDate	Date/Time	PTC - Date Evaluated
PTCStent	Yes/No	PTC - Stent
PTCStentType	Number	PTC - Stent Type

Figure 96: Esophageal Cancer Diagnostic Imaging Table Schema



Figure 97: Esophageal Cancer Preliminary Outlook Form

Field Name	Data Type	Description
80 ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
PreOutlook	Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic or Equivocal Findi

Figure 98: Esophageal Cancer Preliminary Outlook Table Schema

reatment Co	<u>irse</u>	<u>·</u>] ₿ <i>I</i> <u>U</u> ≡ ≡	≡ 2 • Δ • 2 •	
Resection		🖩 Staging Laparos	copy/Laparotomy	1
Radiation		🖩 <u>Palliative Measu</u>	<u>ires</u>	
Adjuvant Neoa	djuvant 🖗 Both	🕅 Bypass	🕅 HAL	
Chemotherapy		🕅 Gastrostomy Tube	🔲 PV Shunts	
Adjuvant R Neoa	djuvant 🖗 Both	🕅 Jejunstomy Tube	🕅 Pall. Stenting	
🕅 Avastin	🕅 Leukovorin	🕅 Celiac Block	🕅 Pall. Radiation	
🕅 Capecitabine	🕅 Levamasole	🕅 Paracentesis	🕅 Pall. Resection	
🕅 Erbitux	🕅 Mitomycin	🕅 Thoracentesis	👿 Other - Specify:	
Fluorouracil (5-FU)	🕅 Oxaliplatin	🕅 Transfusion		
FUDR	🕅 Taxol			
🕅 Gemcitabine	🕅 Other - Specify:		otocol (vaccine, etc.)	
🕅 Irinotecan		Genetic Counsel	ing	

Figure 99: Esophageal Cancer Treatment Form

AutoNumber Text Yes/No Yes/No Yes/No Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	ID Meditech Medical Record Number for Patient Treatment - Resection Treatment - Laparoscopy Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Evolux Treatment - Chemo - Fluorouracil (S-FU) Treatment - Chemo - FUDR Treatment - Chemo - Fuorouracil (S-FU) Treatment - Chemo - Fuorouracil (S-FU)
Yes/No Yes/No Number Yes/No Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Resection Treatment - Laparoscopy Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Evolution Treatment - Chemo - Evolution Treatment - Chemo - Fullor Treatment - Chemo - FUDR Treatment - Chemo - FUDR Treatment - Chemo - Fulcabine
Yes/No Yes/No Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Laparoscopy Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo Treatment - Chemo - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Avastin Treatment - Chemo - Erbitux Treatment - Chemo - Fulorouracii (S-FU) Treatment - Chemo - FUDR Treatment - Chemo - Fulorouracii (S-FU)
Yes/No Number Yes/No Yes/No	Treatment - Radiation Treatment - Radiation - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Avastin Treatment - Chemo - Evolutione Treatment - Chemo - Evolutione Treatment - Chemo - Evolutione Treatment - Chemo - Fluorouracii (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - FUDR Treatment - Chemo - Fuoroania
Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Radiation - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Adjuvancy Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Fuorouracil (5-FU)
Yes/No Number Yes/No	Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gencitabine Treatment - Chemo - FuDR Treatment - Chemo - Fuccan
Number Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Adjuvancy Treatment - Chemo - Avastin Treatment - Chemo - Explicitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracii (S-FU) Treatment - Chemo - FUDR Treatment - Chemo - Fuore Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Avastin Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracii (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Capecitabine Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Erbitux Treatment - Chemo - Fluorouracil (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Fluorouracii (5-FU) Treatment - Chemo - FUDR Treatment - Chemo - Gencitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - FUDR Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No Yes/No	Treatment - Chemo - Gemcitabine Treatment - Chemo - Irinotecan
Yes/No Yes/No Yes/No	Treatment - Chemo - Irinotecan
Yes/No Yes/No	
Yes/No	
	Treatment - Chemo - Leukovorin
	Treatment - Chemo - Levamasole
Testino	Treatment - Chemo - Mitomycin
Yes/No	Treatment - Chemo - Oxaliplatin
Yes/No	Treatment - Chemo - Taxol
Yes/No	Treatment - Chemo - Other
Text	Treatment - Chemo - Other - Specify
Yes/No	Treatment - Palliation
Yes/No	Treatment - Palliation - Pall. Resection
Yes/No	Treatment - Palliation - Bypass
Yes/No	Treatment - Palliation - Celiac Block
	Treatment - Palliation - Paracentesis
	Treatment - Palliation - Thoracentesis
	Treatment - Palliation - Pall, Radiation
	Treatment - Palliation - Transfusion
	Treatment - Palliation - Pall. Stenting
	Treatment - Palliation - PV Shunts
	Treatment - Palliation - HAI
	Treatment - Palliation - Gastrostomy Tube
	Treatment - Palliation - Jejunstomy Tube
	Treatment - Pallation - Other
	Treatment - Palliation - Other - Specify
	Treatment - Experimental protocol (ie. vaccine)
	Treatment - Gene Courseling
	Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No Yes/No

Figure 100: Esophageal Cancer Treatment Table Schema

Microsoft Access - [Eso_6a_Res : Form]
Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5 Image: Serie Format Records Iools Window Help Adobe PDF Type a question for help - 5
If Resection is Performed
Surgery Date of Admission: Date of Admission: Procedure Type If Venous Resection Wenous Resection Other Organs Resected: Estimated Blood Loss (cc): If Transfusion If Yes, Units: Methods: If FFP If Cell Saver
Resection Attempt: © Successful © Unsuccessful - Reason: Post-Op Days in ICU: Post-Op Care Path: © Congruent © Divergent
Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days Image > 7 days
Image: Control Discharge: Discharge Status: Control D
Form View

Figure 101: Esophageal Cancer Resection Form

Field Name	Data Type	Description
(D	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
ResDAdm	Date/Time	Resection - Date of Admission
ResDSurg	Date/Time	Resection - Date of Surgery
ResPxType	Number	Resection - Procedure Type (Whipple, total pancreatectomy, distal pancreatectomy, etc)
ResORTime	Number	Resection - OR Time (hr.)
ResVenRes	Yes/No	Resection - Venous Resection
ResVenRec	Yes/No	Resection - Venous Reconstruction
ResArtRes	Yes/No	Resection - Arterial Resection
ResArtRec	Yes/No	Resection - Arterial Reconstruction
ResOrgans	Text	Resection - Other Organs Resection
ResBloodLoss	Number	Resection - Estimated Blood Loss (cc)
ResTransfusion	Yes/No	Resection - Tranfusion
ResTUnits	Number	Resection - Transfusion Units
ResTFFP	Yes/No	Resection - Transfusion - FFP
ResTCell	Yes/No	Resection - Transfusion - Cell
ResAttempt	Number	Resection - Resection Attempt
ResAttemptUn	Number	Resection - Resection Unsuccessful Reason (Tumor involvement, Operative mishap, etc)
ResPOCourse	Number	Resection - PO - Post-Op Care Path
ResPODays	Number	Resection - PO - Time in ICU (days)
ResPOInfection	Yes/No	Resection - PO - Wound infection
ResPOLeak	Yes/No	Resection - PO - Leak
ResPONG	Yes/No	Resection - PO - NG/gastrotomy drainage
ResPOAbdominal	Yes/No	Resection - PO - Abdominal Collection
ResPOPulmComp	Yes/No	Resection - PO - Pulminary Complications
ResPOLiverInsuf	Yes/No	Resection - PO - Liver Insufficiency
ResPOLiverTB	Number	Resection - PO - Liver Insufficiency - Total Bilirubin
ResPODDischarge	Date/Time	Resection - Date of Discharge
ResPODischStatus	Number	Resection - Discharge Status

Figure 102: Esophageal Cancer Resection Table Schema



Figure 103: Esophageal Cancer No Resection Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
NoResEvalDate	Date/Time	No Resection - Decision Date
NoResNoHandle	Yes/No	No Resection - Couldn't Handle Proposed Treatment
NoResRefused	Yes/No	No Resection - Refused Treatment
NoResMagnitude	Yes/No	No Resection - Magnitude Not Worth Benefits
NoResCeliacInvolve	Yes/No	No Resection - Celiac Trunk Involvement
NoResSMAInvolve	Yes/No	No Resection - SMA Involvement
NoResHepaticInvolve	Yes/No	No Resection - Hepatic Involvement
NoResIVCInvolve	Yes/No	No Resection - Inferior Vena Cava Involvement
NoResSMVInvolve	Yes/No	No Resection - SMV Involvement
NoResPVInvolve	Yes/No	No Resection - Portal Vein Involvement
NoResCirrhosis	Yes/No	No Resection - Cirrhosis
NoResMetastatic	Yes/No	No Resection - Metastatic

Figure 104: Esophageal Cancer No Resection Table Schema



Figure 105: Esophageal Cancer Pathology Form

Field Name	Data Type	Description	
ID MR	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
Histology	Number	Histology	
ResPathT	Number	Resection - Pathology Staging - T	
ResPathN	Number	Resection - Pathology Staging - N	
ResPathM	Number	Resection - Pathology Staging - M	
ResPathR	Number	Resection - Pathology Staging - R	
ResPathV	Number	Resection - Pathology Staging - R	
ResPathSizeX	Number	Resection - Pathology Tumor Size (cm) - Width	
ResPathSizeY	Number	Resection - Pathology Tumor Size (cm) - Height	

Figure 106: Esophageal Cancer Pathology Table Schema

Microsoft Access - [Follow-up Information]	
Eile Edit View Insert Format Records Iools Window Help Adobe PDF Type a question for help	×
K·BQ ⊜Q♥ X B B ∽ % \$ \$ \$ \$ \$ 7 B ∨ \$ H +• K B a• Q.	
Esophageal Tumor	
Follow-up Information Add Record Find Record Delete Record	
MB 000725070 Follow-up Window: 9M V	
Date of Visit: 05/11/2003 Weight (pounds): QOL score:	
ECOG performance status: 6 0 6 1 6 2 6 3 6 4	
Lab Value:	
CEA: Albumin Alkaline Phosphotase	
CA19-9: Total Bilirubin	
Redeveloped Symptoms:	
🗆 Weight loss 👘 Biliary colic 🗖 Pruritis 🗖 Back pain	
how much (pounds)	
☐ Jaundice ☐ Vomiting ☐ Other Specify:	
Cholecystitis Clay colored stool Cholangitis Fatigue	
2 2 2000	
Status:	
© Died Death Date: 5 /11/2003 © N.E.D.	
C A.W.D., Method of Detection: Lab Radiologic Evidence Clinical Evidence	
Form View	11.

Figure 107: Esophageal Cancer Follow-Up Form

12	Field Name	Data Type	Description
I	D	AutoNumber	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
P	1R	Text	Meditech Medical Record Number for Patient
F	UWin	Number	Follow-Up Windows
1	/isitDate	Date/Time	Visit Date
1	Veight	Number	Weight (lbs.)
	OLscore	Number	QoL Score (0-100)
E	COG	Number	ECOG Score (0-4)
L	abCEA	Number	Laboratory - CEA
L	abCA19-9	Number	Laboratory - CA19-9
L	abAlb	Number	Laboratory - Albumin
L	abBili	Number	Laboratory - Bilirubin
L	abAlka	Number	Laboratory - Alkaline phosphotase
19	5xWtloss	Yes/No	Symptoms - Weight Loss
19	5xWtlossP	Number	Symptoms - Weight Loss (lbs.)
19	5xJaun	Yes/No	Symptoms - Jaundice
19	ixChole	Yes/No	Symptoms - Cholecystitis
19	5xChola	Yes/No	Symptoms - Cholangitis
9	5xBC	Yes/No	Symptoms - Biliary Colic
19	5xNau	Yes/No	Symptoms - Nausea
19	5xVom	Yes/No	Symptoms - Vomiting
19	5xCC5	Yes/No	Symptoms - Clay Colored Stool
19	SxFati	Yes/No	Symptoms - Fatigue
19	5xPru	Yes/No	Symptoms - Pruritis
19	ixInd	Yes/No	Symptoms - Indigestion
19	5xAbd	Yes/No	Symptoms - Abdominal Pain
9	5xBack	Yes/No	Symptoms - Back Pain
19	5xOT	Yes/No	Symptoms - Other
19	SXOTSpe	Text	Symptoms - Other - Specify
19	itatus	Number	Status (NED, AWD, Died)
10	DeathDate	Date/Time	Death Date
19	itatusAWDLab	Yes/No	AWD - Lab Evidence
	itatusAWDRad	Yes/No	AWD - Radiology Evidence
	itatusAWDCli	Yes/No	AWD - Clinical Evidence

Figure 108: Esophageal Cancer Follow-Up Table Schema

3.1.6 Colorectal Cancer

- 🖬 🔁 🎒 🗋 💥 🐰 🗉		
	È 🖹 ∽ 🐁 🛃 X↓ 🍞 È ⊽ 🛤 H* M 🖻 🗗 🖢 •	2.
*	- BIUE = A - A - 1 -	+
sumptive Diagnosi	s at Onset of Care	
esentation		
ate of Evaluation ECOG Perfo	ormance Status Height (in.) Weight (lbs.)	
<u>ymptoms</u>		
l Weight Loss	🖩 Bloating 🗰 BRBPR 💭 Constipation	
How Much (pounds):	🕅 Nausea 🛛 🕅 Abdominal Pain 🖉 Hemorrhoids	
l Anal Pain	🖩 Vomiting 🛛 🗐 Back Pain 🖉 Diarrhea	
🕻 Perineal Pain	🖩 Bowel Habit Change 🔎 Other Specify:	
Bowel Obstruction	Time Fatigue	
Doweropstruction		

Figure 109: Colorectal Cancer Presentation Form

	Field Name	Data Type	Description
31	1D	AutoNumber	ID
3	MR	Text	Meditech Medical Record Number for Patient
	PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)
	DemEvalDate	Date/Time	Demographics - Date Evaluated by Surgical Oncology
1	DemECOG	Number	Demographics - ECOG Score (0-4)
	DemHeight	Number	Demographics - Height in Inches of Patient
1	DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission
	SxWtloss	Yes/No	Initial Symptoms - Weight Loss
1	SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds
	SxAnalPain	Yes/No	Initial Symptoms - Anal Pain
	SxPerPain	Yes/No	Initial Symptoms - Perineal Pain
1	SxBObs	Yes/No	Initial Symptoms - Bowel Obstruction
1	SxBloat	Yes/No	Initial Symptoms - Bloating
100	SxNau	Yes/No	Initial Symptoms - Nausea
	SxVom	Yes/No	Initial Symptoms - Vomiting
	SxBHabit	Yes/No	Initial Symptoms - Bowel Habit Change
1	SxFati	Yes/No	Initial Symptoms - Fatigue
-	SxBRBPR	Yes/No	Initial Symptoms - BRBPR
1	SxAbd	Yes/No	Initial Symptoms - Abdominal Pain
	SxBack	Yes/No	Initial Symptoms - Back Pain
1	SxConst	Yes/No	Initial Symptoms - Constipation
	SxHemorr	Yes/No	Initial Symptoms - Hemorrhoids
	SxDiar	Yes/No	Initial Symptoms - Diarrhea
	SXOT	Yes/No	Initial Symptoms - Other
-	SxOTSpe	Text	Initial Symptoms - Other - Specify

Figure 110: Colorectal Cancer Presentation Table Schema

Microsoft Access - [Col_2_History : Form]	<u>_</u> _×
Image: Second secon	
Comorbidities Image: Heart Failure Ischemic Heart Disease Ischemic Heart Disease Respiratory Renal Failure Renal Failure Bleeding Disorder Diet Control Social History Cigarette Use Incohol Use Incohol Use Dither - Specify:	
Form View	

Figure 111: Colorectal Cancer Medical History Form

Field Name	Data Type	Description
ID	AutoNumber	ID
MR	Text	Meditech Medical Record Number for Patient
CxHF	Yes/No	Comorbidities - Heart Failure
CxIHD	Yes/No	Comorbidities - Ischemic Heart Disease
CxResp	Yes/No	Comorbidities - Respiratory
CxDiab	Yes/No	Comorbidities - Diabetes
CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral
CxDiabDiet	Yes/No	Comorbidities - Diabetes - Insulin - Diet Control
CxDiabOnset	Number	Comorbidities - Diabetes - Onset (1=Less than six months, 2 =Greater than six months)
CxRF	Yes/No	Comorbidities - Renal Failure
CxHyper	Yes/No	Comorbidities - Hypertension
CxBleed	Yes/No	Comorbidities - Bleeding Disorder
CxLiver	Yes/No	Comorbidities - Liver Failure
CxMal	Yes/No	Comorbidities - Malnutrition
CxPriorCancer	Number	Comorbidities - Prior Cancer Dx
CxPriorCancerChemo	Yes/No	Comorbidities - Prior Cancer Dx - Chemo
CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation
CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery
SHCigarette	Yes/No	Social History - Cigarettes (significant use)
SHAlcohol	Yes/No	Social History - Alcohol (significant use)
SHDrugÜse	Yes/No	Social History - Drug Use
SHExposure	Yes/No	Social History - Environmental Exposure
SHOther	Yes/No	Social History - Other
SHOtherS	Text	Social History - Other - Specify
FamilyFatherDx	Number	Family History - Father Dx
FamilyMotherDx	Number	Family History - Mother Dx
FamilyOther1	Text	Family History - Other1
FamilyOther1Dx	Number	Family History - Other1 Dx
FamilyOther2	Text	Family History - Other2
FamilyOther2Dx	Number	Family History - Other2 Dx

Figure 112: Colorectal Cancer Medical History Table Schema

Microsoft Acco	ess - [Col_3a_Serum : I	Form]				
B File Edit	View Insert Format	Records Io	ols <u>W</u> indow	Help Adobe F	PDF	- 8
🗹 • 📆	6 D. 🖤 🕺 🖻	B 10 8		V 🖥 🗸 🕅)* W 😭	🗗 📶 🖌 😰 🗸
*	MS Sans Serif	• 8 •	BII	ı ≡≡≡	<u>⊅</u> • <u>∧</u> •	1
erum Stu	dies					
EA:	Albumin	4	ALK	ALT		
A19-9:	Total Bilirubin	4	AST:	Amyla	se:	

Figure 113: Colorectal Cancer Serum Studies Form

Field Name	Data Type	Description	
20 ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
LabCEA	Number	Laboratory - CEA	
LabCA19-9	Number	Laboratory - CA19-9	
LabAlb	Number	Laboratory - Albumin	
LabBili	Number	Laboratory - Bilirubin	
LabAlka	Number	Laboratory - Alkaline phosphotase	
LabALT	Number	Laboratory - ALT	
LabAST	Number	Laboratory - AST	
LabAmylase	Number	Laboratory - Amylase	
LabHPylori	Number	Laboratory - H. Pylori Status	

Figure 114: Colorectal Cancer Serum Studies Table Schema

Microsoft Access - [Col_3b_Diag : Fo	rm]	-OX
E File Edit View Insert Format	Records Tools Window Help Adobe PDF	_ & ×
🔟 - 🖬 😫 🍯 🗟 🖤 🐰 🖻	■ □ ● ● ● ↓ ↓ ▼ ● ▼ ● ▼ ● ● ● ● ● ●	2.
MS Sans Serif	• 8 • B I ∐ ≣ ≣ ≣ ⊉ • ▲ • ⊿ • []•	
Diagnostic and Staging	<u>Studies</u>	-
Colonoscopy		
Last Colonoscopy Date:	Clear of Disease	
Tumor Size (cm):		
🗖 Rectum 🗖 TV Colon	🥅 Hepatic Flexure	
🗖 Sigmoid 🗖 Asc. Colon	🗖 Splenic Flexure	
🗖 Desc. Colon 🗖 Cecum	🗖 Anal Canal	
<u>CT Scan</u>		
CT Scan Clears, Date:		
CT. Scan Shows, Date:	Location 💽	
	How Many Mets:	
Status:	Size of Largest Met (cm):	
		-
Form View		

Figure 115: Colorectal Cancer Diagnostic Imaging Form

ŝ	Field Name	Data Type	Description	
P)	ID	AutoNumber	ID	
	MR	Text	Meditech Medical Record Number for Patient	
Ì	ColLastD	Date/Time	Colonoscopy - Last Date	
	ColClear	Yes/No	Colonoscopy - Clear	
	ColSize	Number	Colonoscopy - Size (cm)	
	ColRectum	Yes/No	Colonoscopy - Rectum	
	ColSigmoid	Yes/No	Colonoscopy - Sigmoid	
Ę	ColDesc	Yes/No	Colonoscopy - Descending Colon	
	ColTV	Yes/No	Colonoscopy - T.V. Colon	
	ColAsc	Yes/No	Colonoscopy - Ascending Colon	
	ColCecum	Yes/No	Colonoscopy - Cecum	
	ColHep	Yes/No	Colonoscopy - Hepatic Flexure	
	ColSplen	Yes/No	Colonoscopy - Splenic Flexure	
	ColAnal	Yes/No	Colonoscopy - Anal Canal	
	CTClearD	Date/Time	CT Scan Clears - Date	
	CTShowD	Date/Time	CT Scan Shows - Date	
	CTLocation	Number	CT Scan Location	
	CTMetCount	Number	CT Met Count	
	CTMetSize	Number	CT Met Size	
	CXR	Yes/No	CXR	
	CXRStatus	Number	CXR	

Figure 116: Colorectal Cancer Diagnostic Imaging Table Schema

Microsoft Access - [Col_4_Prelim : Form]
😰 Eile Edit View Insert Format Records Iools Window Help Adobe PDF Type a question for help 🗸 🗗 🗙
■ ■ ■ ● </th
Pre-Surgical Outlook
Potentially Resectable
Cocally Advanced/Unresectable
Metastatic or Equivocal Findings
Pursue Resectability
🗑 PET 🕅 Special CT Study
Can the Ablation Be Done Percutaneously?
Should Not Reasons:
Could Be Done How: Percutaneous RFA Reasons Open/Laparoscopic
Specialized Techniques Planned
Laparoscopic or Lap Assisted Operation Planned
Open Operation Planned
🖾 Standard Operations Performed 🌾 Resection 🔽
Ablation Resection and/or Ablation
If Not Performed as Planned, Reasons:
If Not Performed as Planned, Result
Planned Attempt Abandoned Completely Planned Attempt Abandoned Extra Attention: Ablation or Resection Too Close, or Positive Margin Resection Added to Operation More Resection Done Resection Added to Operation More Resection Done Planned Pump Not Done
Form View

Figure 117: Colorectal Cancer Preliminary Outlook Form

8	Save Jame	Data Type	Description
P)D	AutoNumber	ID
1	MR	Text	Meditech Medical Record Number for Patient
ŝ	PreOutlook	Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic
	PET	Yes/No	PET Scan
ļ	SpecialCT	Yes/No	Special CT
1	Percut	Number	Percutaneous Ablation -
î	PercutN	Number	Percutaneous Ablation - Should Not - Reasons
È.	PercutY	Number	Percutaneous Ablation - Could Be - How
	PercutYRFA	Number	Percutaneous Ablation - Could Be - RFA
Ì.	Techniq	Number	Specialized Techniques
î	Lap	Number	Laparoscopic Operation Planned
	Operation	Number	Other Operation Planned
ļ	Plan	Yes/No	Standard Operations Performed
Ì.	PlanY	Number	Standard Operations Performed - Types
î	PlanYR	Number	Standard Operations Performed - Resection
	PlanNRes	Number	Standard Operations Not Performed - Reason
	PlanNType	Number	Standard Operations Not Performed - Result
1	PlanNType_PA	Yes/No	Standard Operations Not Performed - Result - Pump Added
î	PlanNType_AAO	Yes/No	Standard Operations Not Performed - Result - Ablation Added to Operation
	PlanNType_RAO	Yes/No	Standard Operations Not Performed - Result - Resection Added to Operation
	PlanNType_OAD	Yes/No	Standard Operations Not Performed - Result - Only Ablation Done
1	PlanNType_ORD	Yes/No	Standard Operations Not Performed - Result - Only Resection Done
î	PlanNType_MR	Yes/No	Standard Operations Not Performed - Result - More Resection Done than Planned
	PlanNType_EA	Yes/No	Standard Operations Not Performed - Result - Extra Attention: Ablation or Resection Too Close
Į.	PlanNType_PND	Yes/No	Standard Operations Not Performed - Result - Pump Not Done

Figure 118: Colorectal Cancer Preliminary Outlook Table Schema

Microsoft Access - [0	[ol_5_Treatment : Form]				
Ele Edit View	Insert Format Records	<u>T</u> ools <u>W</u> indow <u>H</u> elp Ac	dobe PDF Type a question for help	• _ 8 ×	
12 - 2 2 5 0	à 🖤 🐰 🖻 💼 🗠	AI XI ¥ B ∀ A	M >* W @ @ 10 + 0	2).	
*	*	• ₿ ℤ ⊻ 📰 🚍	= <u>A</u> · <u>A</u> · <u>I</u> · <u>-</u> ·		
Treatment Cou	urse				
Resection		Staging Laparos	copy/Laparotomy		
Radiation		🗏 Palliative Measu	Ires		
Adjuvant S Neoad	djuvant 🖗 Both	🐺 Bypass	T HAL		
Chemotherapy		Gastrostomy Tube	₩ PV Shunts		
C Adjuvant C Neoad	djuvant @ Both	🖾 Jejunstomy Tube	🕅 Pall. Stenting		
🕅 Avastin	🕅 Leukovorin	🕅 Celiac Block	🕅 Pall. Radiation		
🕅 Capecitabine	🕅 Levamasole	🕅 Paracentesis	🕅 Pall. Resection		
🕅 Erbitux	🕅 Mitomycin	🕅 Thoracentesis	🔟 Other - Specify:		
Fluorouracil (5-FU)	👿 Oxaliplatin	🕅 Transfusion			
FUDR	🕅 Taxol				
🕅 Gemcitabine	🕅 Other - Specify:	Experimental Protocol (vaccine, etc.)			
🕅 Irinotecan		Genetic Counsel	Genetic Counseling		
Form View				1	

Figure 119: Colorectal Cancer Treatment Form
	Field Name	Data Type	Description
DD		AutoNumber	ID
MR		Text	Meditech Medical Record Number for Patient
TxResect		Yes/No	Treatment - Resection
TxLap		Yes/No	Treatment - Laparoscopy
TxRadia		Yes/No	Treatment - Radiation
TXRadiaA	dju	Number	Treatment - Radiation - Adjuvancy
TxChemo	1 ³⁵¹	Yes/No	Treatment - Chemo
TXChemo	Adju	Number	Treatment - Chemo - Adjuvancy
TxChemo	AVA	Yes/No	Treatment - Chemo - Avastin
TxChemo	Сар	Yes/No	Treatment - Chemo - Capecitabine
TxChemol	Erb	Yes/No	Treatment - Chemo - Erbitux
TxChemol	Flu	Yes/No	Treatment - Chemo - Fluorouracil (5-FU)
TxChemol	FUDR	Yes/No	Treatment - Chemo - FUDR
TxChemo	Gem	Yes/No	Treatment - Chemo - Gemcitabine
TxChemo:	Iri	Yes/No	Treatment - Chemo - Irinotecan
TxChemol	Leu	Yes/No	Treatment - Chemo - Leukovorin
TxChemol	Lev	Yes/No	Treatment - Chemo - Levamasole
TxChemol	Mit	Yes/No	Treatment - Chemo - Mitomycin
TxChemo	Oxa	Yes/No	Treatment - Chemo - Oxaliplatin
TxChemo	Tax	Yes/No	Treatment - Chemo - Taxol
TxChemo	Oth	Yes/No	Treatment - Chemo - Other
TxChemo	OS	Text	Treatment - Chemo - Other - Specify
TxPal		Yes/No	Treatment - Palliation
TxPalRes		Yes/No	Treatment - Palliation - Pall, Resection
TxPalBypa	ass	Yes/No	Treatment - Palliation - Bypass
TxPalCelia		Yes/No	Treatment - Palliation - Celiac Block
TxPalPara	3	Yes/No	Treatment - Palliation - Paracentesis
TxPalTho		Yes/No	Treatment - Palliation - Thoracentesis
TxPalRad		Yes/No	Treatment - Palliation - Pall, Radiation
TxPalTran	ıs	Yes/No	Treatment - Palliation - Transfusion
TxPalSter	15	Yes/No	Treatment - Palliation - Pall. Stenting
TxPalPV		Yes/No	Treatment - Palliation - PV Shunts
TxPalHAL	a 52	Yes/No	Treatment - Palliation - HAL
TxPalGas	Tube	Yes/No	Treatment - Palliation - Gastrostomy Tube
TxPalJeiT	ube	Yes/No	Treatment - Palliation - Jejunstomy Tube
TxPalOth	1	Yes/No	Treatment - Palliation - Other
TxPalOS		Text	Treatment - Palliation - Other - Specify
TXEXD		Yes/No	Treatment - Experimental protocol (ie, vaccine)
TxGene		Yes/No	Treatment - Gene Counseling

Figure 120: Colorectal Cancer Treatment Table Schema





Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
Per_DIH	Number	Days in Hospital	
Per_ICU	Number	Days in ICU	
Per_Trans	Yes/No	Blood Transfusion	
Per_Trans_Units	Number	Units	
Per_chest	Yes/No	Chest Tube	
Per_pneu	Yes/No	Pneumothorax	
Per_emb	Yes/No	Embolization	
Per_sub	Yes/No	Subsequent Drainage	
Per_pul	Yes/No	Pulmonary Complications	
Per_car	Yes/No	Cardiac Complications	
Per_inf	Yes/No	Infection/Abscess	

Figure 122: Colorectal Cancer Ablation Table Schema

Microsoft Access - [Col_7a_Res : Form]
🗉 Eile Edit View Insert Format Records Iools Window Help Adobe PDF Type a question for help 🔹 🗗 🤉
■ · □ · ▲ □ · ▲ □ · ▲ · · · · · · · · · ·
M5 Sans Serif • 8 • B I U ≣ ≣ ≣ 2 • ▲ • 2 • □ • .
f Resection is Performed
Date of Admission: Date of Surgery: OR Time (hr):
Procedure Type Days in ICU:
CVP cm/H20 During Hepatic Transection/Puncture: cm/h20:
Tidal Volume During Hepatic Transection/Puncture (cc):
Other Organs Resected: Estimated Blood Loss (cc):
🕅 Transfusion If yes, units: 🗾 Methods: 🕅 FFP 🕅 Cell Saver
🗑 Tissue Link 🕅 Finger Fracture 🗑 Argon Beam 🖉 Staplers for Structures 👼 RFA
🗮 CVSA 🖉 Clamp/Crush 🕅 Staplers for Parynchyma 🏾 Fibrin Glue 🖉 Cryoablation
🗰 Wound Infection 🛛 🗰 Congruent with Post Op Care Path 🗰 NG/Gastrostomy Drainage > 7 Days
🔟 Leak 🖉 Divergent from Post Op Care Path 🕅 Abdominal Collection
🕅 Pulmonary Complications 🕅 Catheter Infections 🔤 Drains 🕅 Renal Insufficiency
Liver Insufficiency (Total Bilirubin >5 Inpatient) If Yes, Total Bilirubin:
Discharge Status: Date of Discharge/Death:
Form View

Figure 123: Colorectal Cancer Resection Form

	Field Name	Data Type	Description	
31	(D	AutoNumber	ID	
	MR	Text	Meditech Medical Record Number for Patient	
	DAdm	Date/Time	Date of Admission	
	DSurg	Date/Time	Date of Surgery	
	ORTime	Number	OR Time	
	PxType	Number	Procedure Type	
	Organs	Text	Other Organs Resected	
	BloodLoss	Number	Blood Loss (cc)	
	CVP	Yes/No	CVP cm/H20 During Hepatic Transection/Puncture	
	CVPcm	Number	CVP cm/H20	
	Tidalcc	Number	Tidal Volume During Hepatic Transection/Puncture (cc)	
1	Transfusion	Yes/No	Transfusion - Needed?	
	T_Units	Number	Transfusion - Units	
	T_FFP	Yes/No	Transfusion - Fresh Frozen Plasma	
1	T_Cell	Yes/No	Transfusion - Cell Saver	
1	T_Tissue	Yes/No	Transfusion - Tissue Link	
	T_CVSA	Yes/No	Transfusion - CVSA	
1	T Finger	Yes/No	Transfusion - Finger Fracture	
	T_Clamp	Yes/No	Transfusion - Clamp/Crush	
1	T_Argon	Yes/No	Transfusion - Argon Beam	
	T_Pary	Yes/No	Transfusion - Staplers for Parynchyma	
1	T_Struct	Yes/No	Transfusion - Staplers for Structures	
	T_Glue	Yes/No	Transfusion - Fibrin Glue	
1	T_RFA	Yes/No	Transfusion - RFA	
	T_Cry	Yes/No	Transfusion - Cryoablation	
	ICUdays	Number	Days in ICU	
	Infection	Yes/No	PO - Wound Infection	
	Leak	Yes/No	PO - Leak	

Figure 124: Colorectal Cancer Resection Table Schema



Figure 125: Colorectal Cancer No Resection Form

Field Name	Data Type	Description	
ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
NoResEvalDate	Date/Time	No Resection - Decision Date	
NoResNoHandle	Yes/No	No Resection - Couldn't Handle Proposed Treatment	
NoResRefused	Yes/No	No Resection - Refused Treatment	
NoResMagnitude	Yes/No	No Resection - Magnitude Not Worth Benefits	
NoResCeliacInvolve	Yes/No	No Resection - Celiac Trunk Involvement	
NoResSMAInvolve	Yes/No	No Resection - SMA Involvement	
NoResHepaticInvolve	Yes/No	No Resection - Hepatic Involvement	
NoResIVCInvolve	Yes/No	No Resection - Inferior Vena Cava Involvement	
NoResSMVInvolve	Yes/No	No Resection - SMV Involvement	
NoResPVInvolve	Yes/No	No Resection - Portal Vein Involvement	
NoResCirrhosis	Yes/No	No Resection - Cirrhosis	
NoResMetastatic	Yes/No	No Resection - Metastatic	

Figure 126: Colorectal Cancer No Resection Table Schema



Figure 127: Colorectal Cancer Pathology Form

Field Name	Data Type	Description	
)D	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
Histology	Number	Histology	
ResPathT	Number	Resection - Pathology Staging - T	
ResPathN	Number	Resection - Pathology Staging - N	
ResPathM	Number	Resection - Pathology Staging - M	
ResPathR	Number	Resection - Pathology Staging - R	
ResPathV	Number	Resection - Pathology Staging - R	
ResPathSizeX	Number	Resection - Pathology Tumor Size (cm) - Width	
ResPathSizeY	Number	Resection - Pathology Tumor Size (cm) - Height	

Figure 128: Colorectal Cancer Pathology Table Schema

Microsoft Access - [Follow-up Inl	formation]	
	nat Records Tools Window Help Adobe PDF Type a question for help Image: Ima	• - 8 ×
Colorectal Cancer Follow-up Informati	On Add Record Find Record Delete Record	
MR Date of Visit: ECOG performance status: Cab Value:	Follow-up Window: Image: Constraint of the second	
CEA: Albur CA19-9: Total Redeveloped symptoms:	nin Alkaline Phosphotase Bilirubin	
Weight loss how much (pounds) Jaundice Cholecystitis Cholangitis	Biliary colic Pruritis Indigestion Nausea Abdominal pain Vomiting Back pain Clay colored stool Other Fatigue Fatigue	
Status: © Died Death Date: [© N.E.D. © A.W.D., method of detection:	Lab Radiologic evidence Clinical evidence	
Record: 1 1 1	▶* of 1	

Figure 129: Colorectal Cancer Follow-Up Form

12	Field Name	Data Type	Description
I	D	AutoNumber	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
P	1R	Text	Meditech Medical Record Number for Patient
F	UWin	Number	Follow-Up Windows
1	/isitDate	Date/Time	Visit Date
1	Veight	Number	Weight (lbs.)
	OLscore	Number	QoL Score (0-100)
E	COG	Number	ECOG Score (0-4)
L	abCEA	Number	Laboratory - CEA
L	abCA19-9	Number	Laboratory - CA19-9
L	abAlb	Number	Laboratory - Albumin
L	abBili	Number	Laboratory - Bilirubin
L	abAlka	Number	Laboratory - Alkaline phosphotase
19	5xWtloss	Yes/No	Symptoms - Weight Loss
19	5xWtlossP	Number	Symptoms - Weight Loss (lbs.)
19	5xJaun	Yes/No	Symptoms - Jaundice
19	ixChole	Yes/No	Symptoms - Cholecystitis
19	5xChola	Yes/No	Symptoms - Cholangitis
9	5xBC	Yes/No	Symptoms - Biliary Colic
19	5xNau	Yes/No	Symptoms - Nausea
19	5xVom	Yes/No	Symptoms - Vomiting
19	5xCC5	Yes/No	Symptoms - Clay Colored Stool
19	SxFati	Yes/No	Symptoms - Fatigue
19	5xPru	Yes/No	Symptoms - Pruritis
19	ixInd	Yes/No	Symptoms - Indigestion
19	5xAbd	Yes/No	Symptoms - Abdominal Pain
9	5xBack	Yes/No	Symptoms - Back Pain
19	5xOT	Yes/No	Symptoms - Other
19	SXOTSpe	Text	Symptoms - Other - Specify
19	itatus	Number	Status (NED, AWD, Died)
10	DeathDate	Date/Time	Death Date
19	itatusAWDLab	Yes/No	AWD - Lab Evidence
	itatusAWDRad	Yes/No	AWD - Radiology Evidence
	itatusAWDCli	Yes/No	AWD - Clinical Evidence

Figure 130: Colorectal Cancer Follow-Up Table Schema

3.2 Breast Cancer Database

The table schema and interface layout was designed with the help of UMass Medical School oncologists through one-on-one work and efforts of a database committee headed by Dr. Robert Quinlan.

File Edit View Insert Formal	t <u>R</u> ecords <u>T</u> ools <u>W</u> indow <u>H</u> elp Ad	b <u>b</u> e PDF	Type a question for help	7
• 🖬 🗞 🎒 🖧 🖤 🕺 🖻		₩ +* % 😭 🔂 /a • 1 ≡ Δ • Δ • <i>Д</i> • ¯ •		
Breast Cancer S	creening			
Patient Information				<u>19</u>
Demographics:				
MR	Geography			
Height (in.)	Ethnicity			
Aveight (lbs.)	Birthplace			
QOL	Grew Up			
ECOG	Lived Longest			
Pack Years Quit Yr Enviro. Exposure 🕅 Other	rs Ago Drug Use 🔳 Factors 🗐 Other - Specify			
Medical/Family Histo	אופ			
	<u>אור</u>	- Other Cancer		
Medical/Family Histo	OTY Other 1 Breast Dx 🗐	- Other Concer Father Dx		
Medical/Family Histo Breast Cancer Mother Breast Dx 🗐		CONTRACTOR - UNITED AND A PRIME		
Medical/Family Histo Breast Cancer Mother Breast Dx II Mother Breast Dx - Age Sister 1 Breast Dx II	Other 1 Breast Dx Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx	Father Dx	×.	
Medical/Family Histo Breast Cancer Mother Breast Dx Mother Breast Dx - Age Sister 1 Breast Dx - Age Sister 1 Breast Dx - Age	Other 1 Breast Dx 🗐 Other 1 Breast Dx	Father Dx	× . × . × . × .	
Medical/Family Histo Breast Cancer Mother Breast Dx Mother Breast Dx - Age Sister 1 Breast Dx - Age Sister 1 Breast Dx - Age Sister 2 Breast Dx Sister 2 Breast Dx	Other 1 Breast Dx Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx	Father Dx Mother Dx Other Relation 1		
Medical/Family Histo Breast Cancer Mother Breast Dx Mother Breast Dx Sister 1 Breast Dx Sister 1 Breast Dx Sister 1 Breast Dx - Age Sister 2 Breast Dx Sister 2 Breast Dx - Age	Other 1 Breast Dx Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx	Father Dx State Dx St		
Medical/Family Histo Breast Cancer Mother Breast Dx Mother Breast Dx Sister 1 Breast Dx Sister 1 Breast Dx Sister 2 Breast Dx Sister 2 Breast Dx Age Daughter 1 Breast	Other 1 Breast Dx Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx	Father Dx Mother Dx Other Relation 1 Related Dx Other Relation 2		
Medical/Family Histo Breast Cancer Mother Breast Dx Mother Breast Dx Sister 1 Breast Dx Sister 1 Breast Dx Sister 1 Breast Dx - Age Sister 2 Breast Dx Sister 2 Breast Dx - Age	Other 1 Breast Dx Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx	Father Dx Mother Dx Other Relation 1 Related Dx Other Relation 2		

Figure 131: Breast Cancer Screening Form

- R	Field Name	Data Type AutoNumber	Description ID
Ľ	MR	Text	MR .
-	DemMR	Text	Demographics - Meditech Medical Record Number for Patient
	DemHeight DemWeight	Number Number	Demographics - Height in Inches of Patient Demographics - Weight in Pounds of Patient at Admission
	DemQOL	Number	Demographics - QOL Score (0-100)
	DemECOG DemSHCigarette	Number Yes/No	Demographics - ECOG Score (0-4) Demographics - Social History - Cigarettes (significant use)
1	DemSHCigarettePYrs	Number	Demographics - Social History - Cigarettes (significant use) Demographics - Social History - Cigarettes - Pack Years
1	DemSHCigaretteQuit	Yes/No	Demographics - Social History - Cigarettes - Quit?
	DemSHCigaretteQuitYrs DemSHAlcohol	Text Yes/No	Demographics - Social History - Cigarettes - Quit? - Years Demographics - Social History - Alcohol (significant use)
	DemSHDrugUse	Yes/No	Demographics - Social History - Drug Use
0	DemSHExposure	Yes/No	Demographics - Social History - Environmental Exposure
	DemSHOther DemSHOtherS	Yes/No Text	Demographics - Social History - Other Demographics - Social History - Other - Specify
	DemEthnicity	Text	Demographics - Ethnicity
1	DemGeographyBP	Text	Demographics - Geography - Birthplace
	DemGeographyGU DemGeographLL	Text Text	Demographics - Geography - Grew Up Demographics - Geography - Lived Longest
	FamilyBreastMother	Yes/No	Family History - Mother Breast Dx
1	FamilyBreastMotherAge	Number	Family History - Mother Breast Dx Age
	FamilyBreastSister1 FamilyBreastSister1Age	Yes/No Number	Family History - Sister1 Breast Dx Family History - Sister1 Breast Dx Age
Ŭ	FamilyBreastSister2	Yes/No	Family History - Sister2 Breast Dx
1	FamilyBreastSister2Age	Number	Family History - Sister2 Breast Dx Age
	FamilyBreastDaughter1 FamilyBreastDaughter1Age	Yes/No Number	Family History - Daughter1 Breast Dx Family History - Daughter1 Breast Dx Age
Ū	FamilyBreastDaughter2	Yes/No	Family History - Daughter2 Breast Dx
1	FamilyBreastDaughter2Age	Number	Family History - Daughter2 Breast Dx Age
-	FamilyBreastOther1 FamilyBreastOther1Age	Yes/No Number	Family History - Other1 Breast Dx Family History - Other1 Breast Dx Age
	FamilyBreastOther2	Yes/No	Family History - Other2 Breast Dx
	FamilyBreastOther2Age	Number	Family History - Other2 Breast Dx Age
	FamilyFatherDx FamilyMotherDx	Number Number	Family History - Father Dx Family History - Mother Dx
	FamilyOther1	Text	Family History - Other1
	FamilyOther1Dx FamilyOther2	Number Text	Family History - Other1 Dx
-	FamilyOther2 FamilyOther2Dx	Text Number	Family History - Other2 Family History - Other2 Dx
	PriorBiopDate1	Date/Time	Prior Breast Biopsy - Date1
	PriorBiopPlace1	Text	Prior Breast Biopsy - Place1 Prior Breast Biopsy - Type1
	PriorBiopType1 PriorBiopFindings1	Text Text	Prior breast Biopsy - Type1 Prior Breast Biopsy - Findings1
	PriorBiopDate2	Date/Time	Prior Breast Biopsy - Date2
1	PriorBiopPlace2	Text	Prior Breast Biopsy - Place2
	PriorBiopType2 PriorBiopFindings2	Text Text	Prior Breast Biopsy - Type2 Prior Breast Biopsy - Findings2
Ĩ.	ReproMenses	Number	Repro History - Menses Onset
	ReproLastMenstrual	Date/Time	Repro History - Last Menstrual Date Repro History - Menopause Type
0.1	ReproMenoType ReproPregnancies	Number Number	Repro History - Menopause Type Repro History - Pregnancies
j.	ReproChildren	Number	Repro History - Children
1	ReproAbortions ReproAgeFirstBirth	Number Number	Repro History - Abortions
12	ReproAgeMenopause	Number	Repro History - Age of First Birth Repro History - Age of Menopause
	ReproBreastFeeding	Number	Repro History - Breast Feeding
-	ReproHormoneOral ReproHormoneFertility	Number Number	Repro History - Hormone - Oral Contraceptives
8	ReproHormoneHRT	Number	Repro History - Hormone - Fertility Treatments Repro History - Hormone - Hormone Replacement Therapy
	ReproHormoneHolistic	Number	Repro History - Hormone - Holistic/Homeopathic
	ReproGeneticCounseling SxSympPrimary	Memo Number	Repro History - Genetic Counseling Notes
2	SxSympSecondary1	Number	Initial Symptoms - Primary Symptom Initial Symptoms - Secondary Symptom
	SxSympSecondary2	Number	Initial Symptoms - Secondary Symptom Initial Symptoms - Secondary Symptom
15	5x5ymp5econdary3 5xBreastMassDiscovery	Number Text	Initial Symptoms - Secondary Symptom Initial Symptoms - Breast Mass - Discovery
8	SxMastalgia	Yes/No	Initial Symptoms - Discovery Initial Symptoms - Mastalgia
1	SxMastalgiaType	Number	Initial Symptoms - Mastalgia - Type
	SxNDischarge SxNDischargeType	Yes/No Text	Initial Symptoms - Nipple Discharge Initial Symptoms - Nipple Discharge - Type
	SxAxillaryMass	Yes/No	Initial Symptoms - Nilpie Discharge - Type Initial Symptoms - Axillary Mass
2	SxRetraction	Yes/No	Initial Symptoms - Nipple/Skin Retraction
-	SxSystemicPain SxSystemicWeightLoss	Yes/No Yes/No	Initial Symptoms - Systemic - Pain Initial Symptoms - Systemic - Weight Loss
1	SxSystemicWeightLoss SxSystemicWeightLossLbs	Number	Initial Symptoms - Systemic - Weight Loss Initial Symptoms - Systemic - Weight Loss Lbs.
2	SxSystemicDyspnea	Yes/No	Initial Symptoms - Systemic - Dyspnea
-	5x5ystemicNausea 5x5ystemicOTSpe	Yes/No Memo	Initial Symptoms - Systemic - Pain Initial Symptoms - Systemic - Other
	WorkUpScreening	Yes/No	Workup - Screening
-	WorkUpScreeningType	Number	Workup - Screening - Type
-	WorkUpBreastMass WorkUpBreastMassSizeX	Yes/No Number	Workup - Breast Mass Workup - Breast Mass - Size - Width (cm)
	WorkUpBreastMassSizeY	Number	Workup - Breast Mass - Size - Height (cm)
	WorkUpMammo	Number	Workup - Mammogram
-	WorkUpBiRad WorkUpUltra	Number Number	Workup - BIRAD Workup - Ultrasound
	WorkUpMRI	Number	Workup - MRI
2	BiopPalFNA BiopPalFNAImg	Yes/No Yes/No	Biopsy - Palpable - Fine Needle Aspiration Biopsy - Palpable - Fine Needle Aspiration - Imaging
1	BiopPalCNB	Yes/No	Biopsy - Palpable - Core-cutting Needle Biopsy
2	BiopPalCNBImg	Yes/No	Biopsy - Palpable - Core-cutting Needle Biopsy - Imaging
4	BiopPalIncB BiopPalIncBImg	Yes/No Yes/No	Biopsy - Palpable - Incisional Biopsy Biopsy - Palpable - Incisional Biopsy - Imaging
-	BiopPalExcB	Yes/No	Biopsy - Palpable - Excisional Biopsy
2	BiopPalExcBImg	Yes/No	Biopsy - Palpable - Excisional Biopsy - Imaging
4	BiopNonPalMammo BiopNonPalUltra	Yes/No Yes/No	Biopsy - Non-palpable - Mammography Biopsy - Non-palpable - Ultrasound
-	BiopNonPalUltra BiopDisp	Text	Biopsy - Disposition
	CBiopPalFNA	Yes/No	Contra Biopsy - Palpable - Fine Needle Aspiration
4	CBiopPalFNAImg CBiopPalCNB	Yes/No Yes/No	Contra Biopsy - Palpable - Fine Needle Aspiration - Imaging Contra Biopsy - Palpable - Core-cutting Needle Biopsy
-	CBiopPalCNBImg	Yes/No	Contra Biopsy - Palpable - Core-cutting Needle Biopsy Contra Biopsy - Palpable - Core-cutting Needle Biopsy - Imaging
1	CBiopPalIncB	Yes/No	Contra Biopsy - Palpable - Incisional Biopsy
4	CBiopPalIncBImg CBiopPalExcB	Yes/No Yes/No	Contra Biopsy - Palpable - Incisional Biopsy - Imaging Contra Biopsy - Palpable - Excisional Biopsy
-	CBiopPalExcB	Yes/No	Contra Biopsy - Palpable - Excisional Biopsy Contra Biopsy - Palpable - Excisional Biopsy - Imaging
000	CBiopNonPalMammo	Yes/No	Contra Biopsy - Non-palpable - Mammography
-	CBiopNonPalUltra CBiopDisp	Yes/No Text	Contra Biopsy - Non-palpable - Ultrasound Contra Biopsy - Disposition
j.	1		

Figure 132: Breast Cancer Screening Table Schema

crosoft Access - [Breast_Staging] File Edit View Insert Format Recor	ds Tools	Window <u>H</u> elp Ado <u>b</u> e	PDF	Type a guestion for help	
	and the states		.)+∗ ⊯ @ @ (<u>a</u> • [2]		Part 177
- Arial - 9	1.000				
	UNICAL PRACT			<u>-].</u> •	1000
Breast Cancer Stagi	ng				
MR Date of Examination		a a a a a a a a a a a a a a a a a a a			
Physical Examination - Ipsalate	eral				
Breast Mass Size (cm) by		Skin Changes	Nipple Changes		
Clock Dis. from Areola (cm)		A DESCRIPTION OF A DESC	Retraction		
Consistency		Edema 🔳	Discoloration 🔳		
Skin Fixation		Dimpling 🔳	Erosion 🔳		
Pectoral Fixation		Sat. Nodes 🔳	Discharge 🔳		
Chest Wall Fixation		Ulceration 🔳			
Other Features					
Extra Nipples 🔲 Arm Edema 🗏					
Axillary Nodes Node Size (cm)	by	Type			
Supraclav Nodes 🗏 Node Size (cm)	by	Type			
Physical Examination - Contral	lateral				
Breast Mass Size (cm) by	aterai	Skin Changes	Nipple Changes		
Clock Dis. from Areola (cm)		Erythema	Retraction		
Consistency		Edema 🔳	Discoloration		
and the second		Dimpling 🔳	Erosion		
Skin Fixation		Sat. Nodes 🔳	Discharge 🔳		
Skin Fixation Pectoral Fixation		odu nouco —			
		Ulceration			
Pectoral Fixation					
Pectoral Fixation Chest Wall Fixation Other Features	by				

Figure 133: Breast Cancer Staging Form

Field Name	Data Type	Description
ID	AutoNumber	
MR	Text	MR
Date	Date/Time	Date of Examination
Ipsa_Mass_SizeX	Number	Physical Exam - Ipsa - Breast Mass - SizeX
Ipsa_Mass_SizeY	Number	Physical Exam - Ipsa - Breast Mass - SizeY
Ipsa_Mass_Clock	Number	Physical Exam - Ipsa - Breast Mass - Clock
Ipsa_Mass_AreolaD	Number	Physical Exam - Ipsa - Breast Mass - Areola Distance
Ipsa_Mass_Consistency	Text	Physical Exam - Ipsa - Breast Mass - Consistency
Ipsa_Mass_FixSkin	Yes/No	Physical Exam - Ipsa - Breast Mass - Skin Fixation
Ipsa_Mass_FixPect	Yes/No	Physical Exam - Ipsa - Breast Mass - Pectoral Fixation
Ipsa_Mass_FixChest	Yes/No	Physical Exam - Ipsa - Breast Mass - Chest Wall Fixation
[psa_Skin_Erythema	Yes/No	Physical Exam - Ipsa - Skin Changes - Erythema
Ipsa_Skin_Edema	Yes/No	Physical Exam - Ipsa - Skin Changes - Edema
Ipsa_Skin_Dimpling	Yes/No	Physical Exam - Ipsa - Skin Changes - Dimpling
Ipsa Skin Satellite	Yes/No	Physical Exam - Ipsa - Skin Changes - Satellite Nodes
Ipsa_Skin_Ulceration	Yes/No	Physical Exam - Ipsa - Skin Changes - Ulceration
Ipsa Nipple Retraction	Yes/No	Physical Exam - Ipsa - Nipple Changes - Retraction
	Yes/No	
Ipsa_Nipple_Discolor		Physical Exam - Ipsa - Nipple Changes - Discoloration
Ipsa_Nipple_Erosion	Yes/No	Physical Exam - Ipsa - Nipple Changes - Erosion
Ipsa_Nipple_Discharge	Yes/No	Physical Exam - Ipsa - Nipple Changes - Discharge
Ipsa_Extra_Nipples	Yes/No	Physical Exam - Ipsa - Extra Nipples
Ipsa_Nodes_Axil	Yes/No	Physical Exam - Ipsa - Axillary Nodes
Ipsa Nodes Axil SizeX	Number	Physical Exam - Ipsa - Axillary Nodes - SizeX
Ipsa_Nodes_Axil_SizeY	Number	Physical Exam - Ipsa - Axillary Nodes - SizeY
Ipsa_Nodes_Axil_Type	Number	Physical Exam - Ipsa - Axillary Nodes - Type
Ipsa_Nodes_Clav	Yes/No	Physical Exam - Ipsa - Supraclavicular Nodes
Ipsa_Nodes_Clav_SizeX	Number	Physical Exam - Ipsa - Supraclavicular Nodes - SizeX
Ipsa_Nodes_Clav_SizeY	Number	Physical Exam - Ipsa - Supraclavicular Nodes - SizeY
Ipsa_Nodes_Clav_Type	Number	Physical Exam - Ipsa - Supraclavicular Nodes - Type
Ipsa Arm Edema	Yes/No	Physical Exam - Ipsa - Arm Edema
Contra_Mass_SizeX	Number	Physical Exam - Contra - Breast Mass - SizeX
Contra_Mass_SizeY	Number	Physical Exam - Contra - Breast Mass - SizeY
Contra_Mass_Clock	Number	Physical Exam - Contra - Breast Mass - Clock
Contra_Mass_AreolaD	Number	Physical Exam - Contra - Breast Mass - Areola Distance
Contra_Mass_Consistency	Text	Physical Exam - Contra - Breast Mass - Consistency
Contra_Mass_FixSkin	Yes/No	Physical Exam - Contra - Breast Mass - Skin Fixation
Contra Mass FixPect	Yes/No	Physical Exam - Contra - Breast Mass - Pectoral Fixation
Contra_Mass_FixChest	Yes/No	Physical Exam - Contra - Breast Mass - Chest Wall Fixation
Contra_Skin_Erythema	Yes/No	Physical Exam - Contra - Skin Changes - Erythema
Contra_Skin_Edema	Yes/No	Physical Exam - Contra - Skin Changes - Edema
Contra_Skin_Dimpling	Yes/No	Physical Exam - Contra - Skin Changes - Dimpling
Contra_Skin_Satellite	Yes/No	Physical Exam - Contra - Skin Changes - Satellite Nodes
Contra_Skin_Ulceration	Yes/No	Physical Exam - Contra - Skin Changes - Ulceration
Contra_Nipple_Retraction	Yes/No	Physical Exam - Contra - Nipple Changes - Retraction
Contra_Nipple_Discolor	Yes/No	Physical Exam - Contra - Nipple Changes - Discoloration
Contra_Nipple_Erosion	Yes/No	Physical Exam - Contra - Nipple Changes - Erosion
Contra_Nipple_Discharge	Yes/No	Physical Exam - Contra - Nipple Changes - Discharge
Contra_Extra_Nipples	Yes/No	Physical Exam - Contra - Extra Nipples
Contra_Nodes_Axil	Yes/No	Physical Exam - Contra - Axillary Nodes
Contra_Nodes_Axil_SizeX	Number	Physical Exam - Contra - Axillary Nodes - SizeX Physical Exam - Contra - Axillary Nodes - SizeX
Contra_Nodes_Axil_SizeY	Number	Physical Exam - Contra - Axillary Nodes - SizeY
Contra_Nodes_Axil_Type	Number	Physical Exam - Contra - Axillary Nodes - Type
Contra_Nodes_Clav	Yes/No	Physical Exam - Contra - Supraclavicular Nodes
Contra_Nodes_Clav_SizeX	Number	Physical Exam - Contra - Supraclavicular Nodes - SizeX
Contra_Nodes_Clav_SizeY	Number	Physical Exam - Contra - Supraclavicular Nodes - SizeY
Contra_Nodes_Clav_5izer	Number	Physical Exam - Contra - Supraclavicular Nodes - Type
Contra_Arm_Edema	Yes/No	Physical Exam - Contra - Arm Edema
Imaging_Study_1	Number	Imaging - Study - 1
Imaging_Study_1_SizeX	Number	Imaging - Study - 1 - SizeX
Imaging_Study_1_SizeY	Number	Imaging - Study - 1 - SizeY
Imaging_Study_2	Number	Imaging - Study - 2
Imaging_Study_2_SizeX	Number	Imaging - Study - 2 - SizeX
Imaging_Study_2_SizeX	Number	
		Imaging - Study - 2 - SizeY
Imaging_Study_3	Number	Imaging - Study - 3
Imaging_Study_3_SizeX	Number	Imaging - Study - 3 - SizeX
Imaging_Study_3_SizeY	Number	Imaging - Study - 3 - SizeY
Imaging_Study_4	Number	Imaging - Study - 4
Imaging_Study_4_SizeX	Number	Imaging Study - 4 - SizeX
		anoging Jobay TT JAGA
Imaging_Study_4_SizeY	Number	Imaging - Study - 4 - SizeY
Imaging_Study_5	Number	Imaging - Study - 5
Imaging_Study_5_SizeX	Number	Imaging - Study - 5 - SizeX
Imaging_Study_5_SizeY	Number	Imaging - Study - 5 - SizeY
Primary Stage	Number	Staging - Primary Tumor Stage
Node_Stage	Number	Staging - Lymph Node Stage
Mets_Stage	Number Number	Staging - Metastatic Stage Preliminary Outlook
Prelim Outlook		

Figure 134: Breast Cancer Staging Table Schema

MR	Date of Admission				
<u>Operative</u> Breast Side	and a fragment of the second	Resection Type	Success	<u>.</u>	
Surgical I	Reconstruction				
Prosthetic	tions	osthetic Type Tissue Alor	ne Type -		
Bleeding [🛛 Infection 🔳 Seron	na 🔳			
Other Det	ails Muration (hrs				

Figure 135: Breast Cancer Resection Form

Field Name	Data Type	Description
ID	AutoNumber	
MR	Text	MR
Date	Date/Time	Date of Admission/Surgery
Breast_Side	Number	Breast Side
Resection_Type	Number	Resection Type
Reconstruction	Yes/No	Reconstruction
Reconstruction_Type	Number	Reconstruction - Type
Pros_Type	Number	Reconstruction - Prosthetic Type
Tiss_Pros_Type	Number	Reconstruction - Tissue/Prosthetic Type
Tiss_Alone_Type	Number	Reconstruction - Tissue Alone Type
Drain_Use	Yes/No	Drain Use
Drain_Use_Duration	Number	Drain Use - Duration
Comp_Bleed	Yes/No	Complications - Bleeding
Comp_Infection	Yes/No	Complications - Infection
Comp_Seroma	Yes/No	Complications - Seroma
Other_Catheter	Yes/No	Other - Port-a-Cath
Other_Biopsies	Text	Other - Additional Organ Biopsies
Success	Number	Resection Success

Figure 136: Breast Cancer Resection Table Schema

/icrosoft Access - [Breast_Chemo]	
Elle Edit View Insert Format Records Tools Window Help Adobe PDF	Type a question for help 👻 🗕 🛃
· • • • • • • • • • • • • • • • • • • •	
- Arial • 9 • B I U 臣 喜 園 এ • A •	
Breast Cancer Chemotherapy	
MR Date of First Treatment Date of Last Treatment	
Treatment Course	
Chemo Type	
AC M CMF DosaCT M TAC M	
FEC FAC Other Specify	
TEC III VIIIE SPECIAL	
Include Herpacin? 🕅 Neoadjuvant Response Dose Delays	
• 0	
Growth Factors	
Port-a-Cath Placed?	
Complete Therapy?	
Complications	
Dose Dense? Cognitive Changes Hospitalization Febrile Neutro Emesis Cardiac	
Neuropathy Amenorrhea Other Specify	
ard: I∢ + I → I → I → I (Filtered)	
ord: 14 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
	FLTR

Figure 137: Breast Cancer Chemotherapy Form

Field Name	Data Type	Description
D	AutoNumber	
MR	Text	MR
Date_First	Date/Time	Date of First Treatment
Date_Last	Date/Time	Date of Last Treatment
Chemo_Type	Number	Chemo - Type
Chemo_AC	Yes/No	Chemo - AC
Chemo_CMF	Yes/No	Chemo - CMF
Chemo_DosACT	Yes/No	Chemo - DosACT
Chemo_TAC	Yes/No	Chemo - TAC
Chemo_FEC	Yes/No	Chemo - FEC
Chemo_FAC	Yes/No	Chemo - FAC
Chemo_Other	Yes/No	Chemo - Other
Chemo_OtherS	Text	Chemo - Other - Specify
Chemo_Herpacin	Yes/No	Chemo - Herpacin
Neo_Response	Number	Neoadjuvant Response
Dose_Delays	Number	Dose Delays
Growth_Factors	Text	Growth Factors
Dose_Dense	Yes/No	Dose Dense
Comp_Cog	Yes/No	Complications - Cognitive Changes
Comp_Hosp	Yes/No	Complications - Hospitalization
Comp_Feb	Yes/No	Complications - Febrile Neutro
Comp_Eme	Yes/No	Complications - Emesis
Comp_Car	Yes/No	Complications - Cardiac
Comp_Neu	Yes/No	Complications - Neuropathy
Comp_Ame	Yes/No	Complications - Amenorrhea
Comp_Other	Yes/No	Complications - Other
Comp_OtherS	Text	Complications - Other - Specify
Other_Catheter	Yes/No	Other - Port-a-Cath
Complete	Yes/No	Complete Therapy?

Figure 138: Breast Cancer Chemotherapy Table Schema

icrosoft Access - [Breast _. <u>Fi</u> le <u>E</u> dit <u>V</u> iew <u>I</u> nsert		Window Help Adobe PDF	Type a question for help	5
 ✓ ✓ ✓ ✓ Arial 	× 時 記 い 名 2 ・ 9 ・ 1	4 ¥ 𝔅 𝔅 𝔅 𝔅 𝔅 Β Ι Ψ ΕΞΞ Δ·Δ·	© *a • Q . ∡ • □• ⊡• .	
	er Radiother:			
Sites Treated				
Breast Treatment 🗐 Whole Breast Dose	Total # Fractions	Dose/Fraction		
Tumor Bed Boost Dose	Total # Fractions	Dose/Fraction		
<u>Chest Treatment</u> 🔳 Dose	Total # Fractions	Dose/Fraction		
Scar Boost 🔲 Dose	Total # Fractions	Dose/Fraction		
Axilla Treatment III Dose	Total # Fractions	Dose/Fraction		
Supraclavicular Fossa Ti Dose	eatment III Total # Fractions	Dose/Fraction		
rd: 14 4 1	► ► ► ► ► • • • • • • • • • • • • • • •			
			FLTR	

Figure 139: Breast Cancer Radiotherapy Form

Field Name	Data Type	Description
ID.	AutoNumber	
MR	Text	MR
Date_First	Date/Time	Date of First Treatment
Date_Last	Date/Time	Date of Last Treatment
Breast_Treatment	Yes/No	Breast - Treatment
Breast_Dose	Number	Breast - Whole Breast Dose
Breast_Fractions	Number	Breast - Total # Fractions
Breast FracDose	Number	Breast - Dose/Fraction
Breast Boost	Yes/No	Breast - Tumor Bed Boost
Breast Boost Dose	Number	Breast - Tumor Bed Boost - Dose
Breast Boost Fractions	Number	Breast - Tumor Bed Boost - Total # Fractions
Breast Boost FracDose	Number	Breast - Tumor Bed Boost - Dose/Fraction
Chest Treatment	Yes/No	Chest - Treatment
Chest Dose	Number	Chest - Whole Breast Dose
Chest Fractions	Number	Chest - Total # Fractions
Chest FracDose	Number	Chest - Dose/Fraction
Chest Boost	Yes/No	Chest - Scar Boost
Chest Boost Dose	Number	Chest - Scar Boost - Dose
Chest Boost Fractions	Number	Chest - Scar Boost - Total # Fractions
Chest Boost FracDose	Number	Chest - Scar Boost - Dose/Fraction
Axilla Treatment	Yes/No	Axilla - Treatment
Axilla Dose	Number	Axilla - Whole Breast Dose
Axilla Fractions	Number	Axilla - Total # Fractions
Axilla FracDose	Number	Axilla - Dose/Fraction
Supra Treatment	Yes/No	Supraclavicular Fossa - Treatment
Supra Dose	Number	Supraclavicular Fossa - Whole Breast Dose
Supra Fractions	Number	Supraclavicular Fossa - Total # Fractions
Supra FracDose	Number	Supraclavicular Fossa - Dose/Fraction
IntNodes_Treatment	Yes/No	Internal Mammary Nodes - Treatment
IntNodes Dose	Number	Internal Mammary Nodes - Whole Breast Dose
IntNodes Fractions	Number	Internal Mammary Nodes - Total # Fractions
IntNodes_FracDose	Number	Internal Mammary Nodes - Dose/Fraction
Unsched Interrupt	Yes/No	Unscheduled Treatment Interruption
Unsched Interrupt Days	Number	Unscheduled Treatment Interruption - Days
Unsched Interrupt Why	Text	Unscheduled Treatment Interruption - Why?
Complete	Yes/No	Patient Completion
Complete IfNoWhy	Number	Patient Completion - If No Why
Concur_Chemo_TAM	Yes/No	Concurrent TAM or Chemo?
Toxic RT	Yes/No	Toxicity of RT
Toxic RT Grade	Number	Toxicity of RT - Grade
	1.500000000	

Figure 140: Breast Cancer Radiotherapy Table Schema

Microsoft Access - [Breast_Mets]		_ 🗆 🗵
🗃 Elle Edit View Insert Format Records Iools Window Help Adobe PDF	Type a question for help	- 8 ×
⊻·₽₽⊜Q∜∦₽€♡%₽₽♥₩₩₩₽₽₽₽.		
▼ ▼ B I U 臣吾君 <u>3</u> · <u>A</u> · <u>I</u> · □ ·	•	
Breast Cancer Overt Metastatic Treatment		
N		
MR		
Palliative Measures		
Palliative Surgery? Palliative Radiation?		
Palliative Chemotherapy?		
Other Pall, Measures? Specify		
Endocrine Therapy		
Antiestrogens Progestins # <u># Regimens</u>		
Aromatase Inhibitors 🔳 Bisphosphanates 🔳		
Other Specify		
Record: H + 1 > H + of 1 (Filtered)		
MR FLTR		

Figure 141: Breast Cancer Metastatic Treatment Form

Field Name	Data Type	Description
ID	AutoNumber	
MR	Text	MR
Palliative_Surgery	Yes/No	Palliative Surgery?
Palliative_Radiation	Yes/No	Palliative Radiation?
Palliative_Chemo	Yes/No	Palliative Chemotherapy?
Palliative_Chemo_Type	Number	Palliative Chemotherapy? - Type
Palliative_Chemo_Regimens	Number	Palliative Chemotherapy? - # Regimens
Palliative_Other	Yes/No	Palliative Other?
Palliative_OtherS	Text	Palliative Other? - Specify
Endoc_Anti	Yes/No	Endocrine Therapy - Antiestrogens
Endoc_Prog	Yes/No	Endocrine Therapy - Progestins
Endoc_Arom	Yes/No	Endocrine Therapy - Aromatase Inhibitors
Endoc_Other	Yes/No	Endocrine Therapy - Other
Endoc_OtherS	Text	Endocrine Therapy - Other - Specify
Endoc_Regimens	Number	Endocrine Therapy - # Regimens
Bisphos	Yes/No	Bisphosphanates

Figure 142: Breast Cancer Metastatic Treatment Table Schema

1icrosoft Access - [Breast_F	η <u>μ</u>	
Eile Edit Yiew Insert		Type a question for help 📃 🖬
l • 🖬 🗞 🥥 🖉	x @ @ ∽ @ A X Ÿ @ ♥ ₩ ₩ K I	9 6 m · Q.
- Arial	• 9 • B I U 🗏 🖉 • 🛕 •	· 🖉 • 🔲 • 🕞 • .
Breast Cance	the second se	
Di east Cance	Follow-Op	
MR Date of Visit	FU Window	
Performance Status	김 과학 지원 감독 전 감독 전 전 것 같아. 것	
ECOG Score QoL Score		
Serum Markers		
CEA CA1	-3	
Albumin Biliru	bin Alkaline	
	JIII AIKAIIIIE	
Mammographic Findings	Tumor Status	
New Bedeveloped 6		
New/Redeveloped S	Disease in Ipsalateral New Disease in Contralateral	
	Risk Organ Metastatic Symptoms MDS/Acute Leukemia	
Other Specify		
Patient Status		
Status		
Clinical Detection I at Det	the Dedick size (Detection	
Cinical Detection Lab Det	ection Radiological Detectio	
ord: 14 4 1 2	▶ ▶ ▶ of 1 (Filtered)	

Figure 143: Breast Cancer Follow-Up Form

Field Name	Data Type	Description
ID	AutoNumber	
MR	Text	MR
Date	Date/Time	Date of Visit
FU_Window	Number	FU Window
ECOG	Number	ECOG Score
QoL	Number	QoL Score
LabCEA	Number	CEA
LabCA15-3	Number	CA15-3
LabAlb	Number	Albumin
LabBili	Number	Bilirubin
LabAlka	Number	Alkaline
Mammo Find	Number	Mammographic Findings
Tumor_Status	Text	Tumor Status
Tumor_Status_Prim	Yes/No	Tumor Status - New Primary?
Redev_Rec	Yes/No	Redeveloped Symptoms - Recurrent Disease in Ipsalateral Breast
Redev_Contra	Yes/No	Redeveloped Symptoms - New Cancer in Contralateral Breast
Redev_Organ	Yes/No	Redeveloped Symptoms - New Cancer in Other High-Risk Organ
Redev_Mets	Yes/No	Redeveloped Symptoms - Metastatic Symptoms
Redev_Leu	Yes/No	Redeveloped Symptoms - MDS/Acute Leukemia
Redev_Other	Yes/No	Redeveloped Symptoms - Other
Redev_OtherS	Text	Redeveloped Symptoms - Other - Specify
Status	Number	Status
Status_AWD_Lab	Yes/No	Status - AWD - Lab Detection
Status_AWD_Rad	Yes/No	Status - AWD - Radiological Detection
Status_AWD_Cli	Yes/No	Status - AWD - Clinical Detection
Status_Died_Date	Date/Time	Status - Died - Date of Death
Status Died Cause	Text	Status - Died - Cause of Death

Figure 144: Breast Cancer Follow-Up Table Schema

Breast Canc			
MR Pathology	Date		
Specimen Type		Diagnosis	
	•		
Precancerous Spe	cimen Charact	eristics	
Precancerous Type	a an		
	-		
Ductal Carcinoma	Insitu (DCIS) S	pecimen Characteristics	
Histology	Nuclear Grade		
Comedo Necrosis	Size	Multifocality	
	<u> </u>		
# of DCIS Slides	Margin +	Calcification	
		te Francisco de Tales	
Lobular Carcinom	a Insitu (LCIS) S	pecimen Characteristics	
Histology	Size	Necrosis	
HISTORORY			
Margin	Calcification		

Figure 145: Breast Cancer Pathology Form

Field Name	Data Type	Description
)D	AutoNumber	
MR	Text	MR
Date	Date/Time	Pathology Date
Specimen_Type	Number	Specimen Type
Diag_Type	Number	Diagnosis Type
Precan_Type	Number	Precancerous Type
DCIS_Hist	Number	DCIS Histology
DCIS_Grade	Number	DCIS Grade
DCIS_Comedo	Number	DCIS Comedo Necrosis
DCIS_Size	Number	DCIS Size
DCIS_Multifoc	Number	DCIS Multifocality
DCIS Slides	Number	DCIS Slides
DCIS Margin	Number	DCIS Margin
DCIS_Calc	Number	DCIS Calcification
LCIS Hist	Number	LCIS Histology
LCIS Size	Number	LCIS Size
LCIS Necrosis	Number	LCIS Necrosis
LCIS Margin	Number	LCIS Margin
LCIS Calc	Number	LCIS Calcification
DIC Type	Number	DIC Type
LIC Type	Number	LIC Type
LIC Grade	Number	LIC Grade
LIC Size	Number	LIC Size
LIC LVI	Number	LIC Lymphovascular Invasion
LIC Necrosis	Number	LIC Necrosis
LIC Margin	Number	LIC Margin
LIC Skin	Number	LIC Skin Involvement
LIC Nipple	Number	LIC Nipple Involvement
LIC IS Presense	Number	LIC IS Presense
LIC IS Type	Number	LIC IS Type
LIC IS Grade	Number	LIC IS Grade
LIC IS Invasion	Number	LIC IS Invasion
LIC IS EIC	Number	LIC IS EIC
LIC_Node_Type	Number	LIC Node Type
LIC Node Pos	Number	LIC Node Positivity
LIC_Node_Extracap	Number	LIC Node Extracapsular Invasion
LIC Micro	Number	LIC Microcalcification
LIC_Rec_Estro	Number	LIC Estrogen Receptor
LIC_Rec_Prog	Number	LIC Progesterone Receptor
LIC_Rec_IHC	Number	LIC IHC Receptor
LIC_Rec_FISH	Number	LIC FISH Receptor
LIC_Stage	Number	LIC Pathology Stage

Figure 146: Breast Cancer Pathology Table Schema

4 Clinical Performance Machine Learning -Procedure & Design

4.1 Objectives of Analysis

As the pancreatic cancer module was the most developed and populated module within our database, it was chosen to be the focus of our machine learning analysis. Given the aggressive nature of these tumors, treatment decisions may often be a complex and ambiguous task, particularly in regard to resective surgery. Physicians seek *prediction models* to aid in the application of pancreatic cancer therapies in a clinical setting. Prediction models for pancreatic cancer clinical factors, particularly survival rates, have been suggested based on such factors as TNM staging, age, gender, presentation symptoms, medical comorbidities, tumor histology, and relation of disease to vasculature. The majority of these predictive models in modern oncology literature are generated by *regression algorithms* (e.g. linear regression, logistic regression, and Cox's proportional hazard model) [Tse04, FS03, SR02].

We have chosen a set of *prediction targets* for which to develop prediction models. We use linear and logistic regression algorithms, as well as machine learning *classification algorithms* (Bayesian methods, decision trees, k-nearest-neighbor, multi-layer perceptrons, etc.), to generate prediction models which are novel to pancreatic cancer research. Our hope is that these novel prediction models may enlightened and improve upon current treatment methods. For the preparation and analysis of our data, *pre-processing algorithms* will be used, including supervised discretization and correlation-based feature selection. *Meta-learning algorithms*, such as Bagging and AdaBoostM1, will be used to boost prediction model effectiveness. The accuracy of these novel prediction models will be statistically compared to models generated by traditional regression methods. The prediction targets studied will include tumor size, T-staging, N-staging, vasculature involvement, tumor histology, malignancy, survival rates, and ECOG scores at 6-month, 9-month, and 12-month follow-up intervals.

4.2 Patient Data Set

Our study population is composed of pancreatic cancer patients seen over the past three years at UMass Memorial hospital in Worcester, Massachusetts. Complete screening, treatment, and follow-up records were retrospectively compiled from the hospital's Meditech electronic record system into our clinical database. Supervision by the medical staff was provided for the interpretation of ambiguous or incomplete records. A total of 91 evaluations for pancreatic cancer treatment were done between April 2003 and May 2006, representing 87 unique patients.

During these evaluations, all patients were screened for tumor resection using diagnostic imaging and clinical evaluation. A total of 74 (81%) resections were subsequently performed with a surgical success rate (complete excision of tumor) of 96%. Radiotherapy was assigned in 37 (41%) evaluations, chemotherapy in 39 (43%) evaluations, and palliative measures in 11 (12%) evaluations. Among the tumors evaluated, 75 (82%) were deemed potentially resectable, 7 (8%) locally advanced/unresectable, and 9 (10%) metastatic or equivocal. Patient age at time of enrollment ranged from 28.5 to 85.1, with an average age of 63.9. Among the patients, 49 (56%) were female. Distribution and availability of this study's prediction targets are detailed in Tables 6 through 15.

Our objective of effective data mining was challenged by various aspects of this data set. Only a relatively small number of patient instances were available for the study, which is a frequent concern in oncology research. Studies are often constrained by the number of patients seen at an institution, or the rarity of certain disease etiologies [KBK⁺97]. However, the number of patients available here has proved sufficient in other pancreatic cancer studies [DD04, SR02]. The limited number of patients is made more difficult by the inconsistent availability of certain prediction targets. Factors such as T-stage, N-stage, tumor size, and follow-up ECOG scores are not provided for all patients. Unavailability of clinical factors also extends to many patient attributes.

In an effort to create a detailed clinical database, patient representations in table schemata

are highly dimensional. After serializing attributes are removed, approximately 190 columns of data are processed for each patient instance. Although this creates a very detailed clinical representation of the patient, the attributes vary greatly in importance, accuracy, and availability, which in turn impacts predictive model accuracy. Data typing also varies—both nominal and numeric attributes are captured in a patient instance. As many aspects of the clinical narrative are tracked, from presentation to treatment to follow-up, there are even some theoretical questions as to whether a collaborative interpretation of these factors may be the correct approach.

Finally, there is the issue of skewed class distribution in data sets. In pancreatic cancer, certain values may frequently dominate various clinical factors. For example, in our patient data set, a large majority of the histologic types are ductal adenocarcinoma, T3 value accounts for 76% of all T-stagings, 82% tumors behave in a malignant fashion, and the majority of patients do not require a vascular resection. These data patterns lend themselves to predictive models which underemphasize the importance of correctly predicting non-majority class values.

In our experimental design, various data mining methods are incorporated to compensate for these issues. Use of meta-learning algorithms helps compensate for small data sets and reduces the effect of over-fitting. Supervised discretization creates a uniformly typed set of attributes. Feature selection algorithms pare highly dimensional groups of attributes to smaller sets of independently behaving features which are highly correlated to the target class. Future research will incorporate over-sampling techniques to improve models based on skewed data sets. These techniques will be discussed more thoroughly in the following section.

Value	Count
0.0 - 2.0 cm	19
2.0 - 3.2 cm	20
3.2 - 4.8 cm	18
4.8 cm - inf 17	
N = 74	

Table 6: Tumor Size Distribution

Value	Count		
T0	1		
T1	2		
Τ2	3		
Τ3	T3 39		
Τ4	6		
N = 51			

Table 7: T-Stage Distribution

Value	Count
N0	16
N1	34
N2	1
N = 51	

Table 8: N-Stage Distribution

Value	Count
True	13
False	61
N = 74	

 Table 9: Vasculature Involvement Distribution

Value	Count
Adenocarcinoma of Pancreas - NOS	24
Ampullary Adenocarcinoma	9
Benign Cyst	1
Cystadenoma	4
Distal Cholangiocarcinoma	1
Duodenal Adenocarcinoma	2
Ductal Adenocarcinoma of Pancreas	27
IPMN - Benign or CiS	11
MEN-I	1
Mucinous Cystic Neoplasm	1
Neuroendocrine	5
Pseudopapillary Tumor	1
Renal Mets	3
Von Hippel-Lindau Syndrome	1
N = 91	

Table 10: Histology Distribution

Value	Count
Benign	16
Malignant 75	
N = 91	

Table 11: Malignancy Distribution

Value	Count
0	37
1	27
2	8
N = 68	

Table 12: ECOG 6-Month Distribution

Value	Count
0	33
1	13
2	7
3 4	
N = 57	

Table 13: ECOG 9-Month Distribution

Value	Count
0	23
1	12
2	7
3 2	
N = 34	

Table 14: ECOG 12-Month Distribution

Value	Count
0 - 6 mo.	20
6 - 12 mo.	20
12 - inf mo. 20	
N = 60	

Table 15: Survival Distribution

4.3 Data Mining and Machine Learning Algorithms Used

The following machine learning algorithms are used in our experiments to generate prediction models. In creating prediction models, a target may be interpreted as a *nominal* (categorical) or *numeric* class. The interpretation of the prediction target influence what machine learning algorithms may be applied. Brief descriptions and research citations are provided. All algorithm executions are run using the Weka machine learning workbench [IW05]. The debug parameter is set to False for all algorithm executions.

4.3.1 Benchmark Algorithms

These algorithms generate prediction models which are used as performance benchmarks for our remaining experiments.

- ZeroR Rudimentary zero-knowledge algorithm used to predict entity classification. ZeroR models in nominal prediction choose the most frequently occurring target classification across all available instances. ZeroR models in numeric prediction choose the average target value of available instances [Mit97].
- Linear Regression Algorithm which expresses a numeric class as a linear combination of weighted attributes. The weights of each attribute are calculated based on the training data. Weights are chosen during model generation such that sum of squares of differences between the training and prediction instances is minimized. Weka's implementation of linear regression uses Akaike criterion for model selection. Weka parameters used are attributeSelectionMethod = M5 method, eliminateColinearAttributes = True, ridge = 1.0E-8 [Aka74, Dev95].
- Logistic Regression Works in a similar fashion to linear regression in combining a weighted set of attributes. Used for nominal targets. For dual-class targets, the linear model is based on a logit transformation of the target class. Multiple classes ar generated using pairwise classification. Attribute weights are assigned by maximizing

log-likelihood of the predictive model. Weka parameters used are maxIts = -1, ridge = 1.0E-8 [lCvH92].

4.3.2 Classification Algorithms

Classification algorithms are used to generate prediction models for nominal targets and binned ranges of numeric targets.

- OneR Rudimentary algorithm which uses single-attribute models to predict entity classification. Also known as 1R or Learn-One-Rule. OneR is known for reasonable accuracy in characterizing experimental data in spite of its relative simplicity. Weka parameters used are minBucketSize = 6 [Mit97].
- J48 A Java implementation of the C4.5 decision tree learning algorithm. C4.5 is an evolution of the basic ID3 decision tree algorithm which accounts for missing values, continuous attributes, pruning of decision trees, and rule derivation. Weka parameters used are binarySplits = False, confidenceFactor = 0.25, minNumObj = 2, numFolds = 3, reducedErrorPruning = False, saveInstanceData = False, seed = 1, subtreeRaising = True, unpruned = False, useLaplace = False [IW05, Qui93].
- Locally Weighted Learning Instance-based prediction model which weights training instances in relation to their distance to the test instance. Closer instances are assigned higher weight and more relevance to the prediction. Can be combined with most classifier algorithms. Locally weighted learning plus Naive Bayes is known to be very effective on small data sets and can outperform independent executions of Naive Bayes and k-nearest-neighbor. Weka parameters used are KNN = -1, classifier = NaiveBayes, dontNormalize = False, weightingKernel = 0 [FHP03, AMS97].
- K-Nearest-Neighbor An instance-based model which produces a classification by calculating the k-closest known members in instance space. Assumes attributes are equally important and normalized. Space between attribute values is calculated using

Euclidean distance. Value of k is determined by cross-validation. Weka parameters used are KNN = varies by experiment, crossValidate = False, distanceWeighting = No distance weighting, meanSquared = False, noNormalization = False, windowSize = 0 [AKA91].

- Naive Bayes The NaiveBayes algorithm is a predictive classifier based on probability models rooted in Bayes Theorem. It assumes statistical independence amongst the attributes in predicting a target classification. NaiveBayes offers surprising accuracy in characterizing data from a variety of domains despite its statistical simplicity. Weka parameters used are useKernelEstimator = False, useSupervisedDiscretization = False [Mit97].
- Bayes Net Bayesian networks are directed acyclic graphs which represent complex statistical relationships for attributes of an entity. Bayesian net predictors construct a graph probability model for classification using a specified network evaluator and network-space search function. Weka parameters used are BIFFile = null, estimator = SimpleEstimator -A 0.5, searchAlgorithm = K2 -P [varies by experiment], useADTree = False [IW05].

4.3.3 Regression Algorithms

Regression algorithms are used to generate prediction models for numeric classes.

M5P - A Java implementation of the M5 algorithm. M5 is a decision tree predictor which builds model trees based on information gain measures. These model trees split the data into test outcomes, which are used to produce a set of multivariate linear regression models. Weka allows both regression trees and model trees to be produced as output. Weka parameters used are buildRegressionTree = False, minNumInstances = 4.0, saveInstances = False, unpruned = False, useUnsmoothed = False [IW05, Qui92].

- Multi-layer Perceptron A neural network which uses backpropagation to train network connection weights. The number of layers for each model are determined during the experiment. Attributes and numeric classes are normalized during execution. Weka parameters used are GUI = false, autoBuild = False, decay = False, hiddenLayers = varies by experiment, learningRate = 0.3, momentum = 0.2, nominalToBinaryFilter = True, normalizeAttributes = True, normalizeNumericClass = True, randomSeed = 0, reset = True, trainingTime = 500, validationSetSize = 0, validationThreshold = 20 [IW05].
- Radial Basis Function Network A variation on the multi-layer perceptron which is implemented by a feedforward network. Computation at each hidden node is performed using k-means computation of distance space. The output, or activation, of the node depends on its distance from the input instance-closer distance generates stronger activation. Similarity measures are calculated using a Gaussian activation function. Network output is a linear combination of hidden node outputs. Weka parameters used are clusteringSeed = 1, maxIts = -1, minStdDev = 0.1, numClusters = 2, ridge = 1.0E-8 [MD89].

4.3.4 Data Preprocessing Algorithms

Data preprocessing methods allow us to achieve various representations of the clinical patient data when conducting experiments. These can potentially improve accuracy of the prediction models generated.

• Discretization - Numeric attribute data may be discretized to form nominal attributes. Discretization is either a *supervised* or *unsupervised* process. Unsupervised discretization proceeds by simply binning data into specified ranges. Supervised discretization bins attributes relative to changes in the target classification. Here, we measure changes in target classification using the Minimum Descriptive Length (MDL) principle. Weka parameters used for supervised discretization are attributeIndices = first-last, invert-Selection = False, makeBinary = False, useBetterEncoding = False, useKononenko = False [FI93].

• Feature Selection - Correlation-based Feature Selection (CFS) is an attribute-selection algorithm used for eliminating noisy and redundant features in data sets. Attributes are selected using heuristic search of correlation measurements. Optimal attribute sets exhibit high correlation to their target class and low correlation to other attributes. Feature selection is useful for paring down high-dimensional data. Weka parameters used are evaluator = CfsSubsetEval, search = BestFirst -D 1 -N 5 [Hal98].

4.3.5 Meta-Learning Algorithm

Meta-learning algorithms are used to improve the accuracy of our machine learning tests. Meta-learning refines models to be more robust against noisy data and less susceptible to over-fitting, particularly when dealing with small data sets.

- AdaBoostM1 AdaBoostM1 works by incrementally running classifiers on samples of test data and combining them into an aggregate model. Each individual or *weak* classifier contributes to the aggregate model in proportion to its accuracy. After each iteration, test data is reweighted based on incorrect aggregate classifications. This boosts the emphasis of misclassified instances, which refines future weak classifier executions. Weka parameters used are classifier = varies by experiment, numIterations = 10, seed = 1, useResampling = False, weightThreshold = 100 [FS96].
- Bagging Bagging (or Bootstrap Aggregating) works similarly to Boosting by combining the results of multiple classifiers into an aggregate model. Multiple prediction models are trained and aggregated using equal-sized resamples from the training data. Bagging is known to be particularly useful when small changes in data can imply large

changes in classification. Weka parameters used are bagSizePercent = 100, calcOutOf-Bag = False, classifier = varies by experiment, numIterations = 10, seed = 1 [Bre96].

• Stacking - The Stacking algorithm is a meta-learner which reduces individual bias by combining multiple classifier types. First, a series of general classifiers generate *level-0* prediction models from a given test set. Data assembled from the output of these models is combined by another classifier to generate a *level-1* prediction model. Weka parameters used are classifiers = varies by experiment, metaClassifier = DecisionS-tump, numFolds = 10, seed = 1 [Wol90].

4.4 Experimental Design

Clinical prediction models are generated using classification for nominal targets and regression for numeric targets. The experiment names of nominal targets (which also include binned numeric ranges) are listed in Table 16. The experiment names of numeric targets are listed in Table 17.

Each experiment is performed using 10-fold cross-validation. As some of these experiments are probabilistic in nature, they are repeated over 10 iterations with random seeding. Performance of classification models are evaluated by calculating the average *accuracy* (percentage correct) classifications across these iterations. Regression models are evaluated by calculating *r*-squared values (Equation 1), which define percentage of response variability accounted for by the prediction model [Dev95].

$$r^2 = \frac{ESS}{TSS} \tag{1}$$

ESS stands for *Explained Sum of Squares* (Equation 2). It stands for the sum of squares of the differences of the predicted independent variable (\hat{y}_i) within the regression model and the overall average of actual independent variables, or grand mean (\bar{y}) . TSS stands for *Total Sum of Squares* (Equation 3). It stands for the sum of squares of the differences of the actual independent variable (y_i) and the grand mean.

$$ESS = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$
(2)

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y})^2$$
(3)

Experiment	Prediction Target
C1	Tumor Size (binned)
C2	T-Stage
C3	N-Stage
C4	Vasculature Involvement
C5	Histology
C6	Malignancy
C7	ECOG 6-Month
C8	ECOG 9-Month
C9	ECOG 12-Month
C10	Survival (binned)

 Table 16: Classification Experiments

Experiment	Prediction Target
R1	Tumor Size
R2	ECOG 6-Month
R3	ECOG 9-Month
R4	ECOG 12-Month
R5	Survival

Table 17: Regression Experiments	Table 17:	Regression	Experiments
----------------------------------	-----------	------------	-------------

Category	Symbol	Algorithm
Rule-based	ZR	ZeroR
	1R	OneR
Decision Trees	J48	C4.5 Decision Trees
Lazy Evaluators	IB1	K-Nearest-Neighbor $k=1$
	IB2	K-Nearest-Neighbor $k=2$
	IB3	K-Nearest-Neighbor $k=3$
	LWL	Locally Weighted Learning w/ Naive Bayes
Bayesian Methods	BN1	Bayes Net p=1
	BN2	Bayes Net $p=2$
	BN3	Bayes Net $p=3$
	NVB	Naive Bayes
Regression	LGR	Logistic Regression

Table 18: Classification Algorithms

4.4.1 Classification Tests

The classification algorithms used and their associated parameters are described in Table 18. Each classification algorithm was repeated using AdaBoostM1 (AB1) and Bagging (BG) meta-learners.

Four data sets (A-D) based on each prediction target (C1-C10) were created from the clinical database. Each data set was first anonymized and stripped of serializing attributes (date of admission, medical record number, etc.). Numeric targets (tumor size, survival, etc.) were binned into equal frequency numeric ranges so to be compatible with nominal classification. Classification target ranges, including numeric bins, are described in Table 19. Preprocessing methods were applied to each data set as described in Table 20. Supervised discretization was used to create uniform nominal attributes, which occasionally produces more accurate experimental results [IW05]. Attribute selection was used to pare down the high dimensionality of the original data sets. Frequently, attribute selection produces more accurate prediction models. It was also useful in generating a medically novel set of highly-correlated, independently behaving attributes for the clinical factor in question.

Clinical Factors - Nominal Categories		
Tumor Size	0 - 2.0 cm, 2.0 - 3.2 cm, 3.2 - 4.8 cm, 4.8 cm - inf	
T-Stage	TX - T4	
N-Stage	NX - N2	
Vasculature Involvement	Yes, No	
	Adenocarcinoma of Pancreas - NOS,	
	Ampullary Adenocarcinoma, Ductal Adeno of Pancreas,	
Histology	Neuroendocrine, Duodenal Adenocarcinoma, Distal	
	Cholangiocarcinoma, Renal Mets, Cystadenoma,	
	IPMN - Benign or CiS, Benign Cyst	
Malignancy	Malignant, Benign	
ECOG 6-Month	0 - 4 (Ref. Table 2)	
ECOG 9-Month	0 - 4 (Ref. Table 2)	
ECOG 12-Month	0 - 4 (Ref. Table 2)	
Survival Rate	0 - 7.0 mo., 7.0 - 16.8 mo., 16.8 - inf	

Table 19: Classification Target Values

Data Set	Pre-processing Filters (ref. Section 4.3.4)
	Class Discretization: Discrete target classes are
А	required for classification algorithms. Nominal target classes
	are naturally discrete. Numeric target are discretized via
	unsupervised equal-frequency binning.
В	Supervised Attributes Discretization: Instance attributes
D	are discretized via MDL method. Derived from Data Set A.
С	Correlation-based Feature Selection: Attribute subsets are chosen
C	based on the CFS method. Derived from Data Set A.
D	Correlation-based Feature Selection and Supervised Discretization:
	Uses both MDL discretization and CFS attribute
	selection. Derived from Data Set B.

Table 20: Classification Data Sets
Category	Symbol	Algorithm
Rule-based	ZR	ZeroR
Decision Trees	M5M	M5P w/ Model Trees
Decision frees	M5R	M5P w/ Regression Trees
Neural Network	MLP	Multi-layer Perceptron
neural network	RBF	Radial Basis Function
Regression	LNG	Linear Regression

Table 21: Regression Algorithms

Clinical Factors - I	Numeric Ranges
Tumor Size	0.0 - 11.0 cm
ECOG 6-Month	0 - 2
ECOG 9-Month	0 - 3
ECOG 12-Month	0 - 3
Survival Rate	1.4 - 44.2 mo.

Table 22: Regression Experiments

4.4.2 Regression Tests

The regression algorithms used and their associated parameters are described in Table 21. Regression target numeric ranges are described in Table 22. Each regression run is repeated using Bagging (BG) meta-learners (AdaBoostM1 is unable to handle numeric targets). Additionally, the Stacking (STK) meta-learner is used to combine the M5P decision trees, RBF networks and linear regression models.

Two data sets (E-F) based on each prediction target (R1-R5) were created from the clinical database. Data sets were anonymized and serializing attributes removed as with classification tests. Attribute selection preprocessing methods were applied as described in Table 23. Supervised discretization filtering was not applied as it requires a nominal target class [FI93].

Data Set	Pre-processing Filters (ref. Section 4.3.4)
E	Unaltered Data Set: Uses original
Ľ	instance data with numeric target classes.
Б	Correlation-based Feature Selection: Attribute subsets are chosen
T,	based on the CFS method. Derived from Data Set E.

Table 23: Regression Data Sets

5 Clinical Performance Machine Learning -

Results & Analysis

For each experiment, we present result sets and graphs for basic algorithm executions and executions using meta-learners. For classification tests, we conduct t-tests of performance of algorithms versus logistic regression. For regression tests, t-tests are performed of algorithm performance versus linear regression. All t-tests are performed with significance $\alpha = .05$ [Dev95]. T-test results are denoted with '=' for statistically equivalent performance, '+' for superior performance, and '-' for inferior performance.

5.1 C1 - Tumor Size

For the tumor size tests among N=74 patients, we predict tumor size of 4 numeric bins which contain roughly equal numbers of patients. Distribute of target values is shown in Table 6. Classification accuracy for tumor size prediction generally ranges from 40% to 55%. The majority of algorithms performed comparably to logistic regression via t-testing. Data sets with supervised discretization and attribute selection generally produced more accurate results. No statistically significant change was seen when meta-learning was introduced.

Classif	icatio	n - Tu	mor S	Size								
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	44.96	25.71	46.54	27.23	32.36	33.52	34.48	36.50	36.88	39.25	37.43	34.36
В	39.89	25.71	38.54	29.11	36.43	39.91	42.82	48.45	48.02	42.63	41.52	39.77
С	49.66	25.71	48.36	35.54	50.89	44.80	44.30	54.52	54.61	54.77	50.62	48.21
D	48.77	25.71	38.29	40.59	48.73	45.07	46.16	49.50	54.12	56.32	53.61	52.02

Figure 147: Tumor Size - Accuracy Results (Percentage)

Classif	icatior	ı - Tun	nor Siz	e - Ada	Boost	M1					
Data Set	LGR	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	44.96	35.64	28.73	32.36	31.32	34.63	36.16	35.59	39.29	38.36	38.27
В	39.89	34.54	28.68	36.43	38.20	42.64	48.66	38.71	42.57	40.39	40.00
С	49.66	39.70	39.61	50.89	49.95	45.29	50.66	54.32	50.18	46.45	46.52
D	48.77	33.68	43.05	50.38	46.23	45.54	46.64	50.21	54.20	51.55	50.79

Figure 148: Tumor Size - Accuracy Results (Percentage) - AdaBoostM1

Classif	ication -	- Tum	or Siz	e - Bag	gging						
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	44.96	42.25	30.71	32.50	31.79	33.55	36.86	38.38	42.30	38.96	37.27
В	39.89	33.46	31.84	35.16	38.41	37.73	46.43	45.39	42.20	40.73	38.27
С	49.66	43.25	40.79	49.82	46.82	46.54	54.73	51.91	54.61	49.55	50.16
D	48.77	37.21	40.93	48.27	47.07	46.77	50.50	53.52	54.95	48.54	47.80

Figure 149: Tumor Size - Accuracy Results (Percentage) - Bagging



Figure 150: Tumor Size - Results Graph



Figure 151: Tumor Size - Results Graph - AdaBoostM1



Figure 152: Tumor Size - Results Graph - Bagging

Tumor	Tumor Size T-Test - ML Algorithms vs. Logistic Regression														
Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3				
A	-	=	-	-	=	=	=	=	=	=	=				
В	-	=	=	=	=	=	=	=	=	=	=				
С	-	=	=	=	=	=	=	=	=	=	=				
D	-	=	=	=	=	=	=	=	=	=	=				
+ : Superio	+ : Superior to LGR = : Equivalent to LGR - : Inferior to LGR														

Figure 153: Tumor Size - T-Test vs. Logistic Regression

Tumor	Tumor Size T-Test - ML Algs. w/ AdaBoostM1 vs. Logistic Regression													
Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1				
A	=	-	-	-	=	=	=	=	Ξ	=				
В	=	=	=	=	=	=	=	=	=	=				
С	=	=	=	=	=	=	=	=	=	=				
D	=	=	=	=	=	=	=	=	=	=				

Figure 154: Tumor Size - T-Test vs. Logistic Regression - AdaBoostM1

Data Set	1R BG	.148 BG	IB1 BG	IB2 BG	IB3 BG	I WI BO	NVB BG	BN1 BG	BN2 BG	BN3 BG
A	=	-	-	-	=	=	=	=	=	=
В	=	=	=	=	=	=	=	=	=	¥)
С	=	=	=	=	=	=	=	=	=	=
D	=	=	=	-	-	=	=		=	=

Figure 155: Tumor Size - T-Test vs. Logistic Regression - Bagging

5.2 C2 - T-Stage

For the t-staging tests among N=51 patients, we predict t-stage of 5 classes which are dominated by value T3 (approx. 75% of patients). Distribute of target values is shown in Table 7. Classification accuracy for t-size prediction generally ranges from 70% to 80%. Unfortunately, analysis of the associated confusion matrices show that prediction dominates for the majority T3 class and under-predicts the remaining values. The majority of algorithms in A and B data sets performed better than logistic regression via t-testing-this seems due more to logistic regression's unusually poor performance for these sets. Data sets with supervised discretization and attribute selection generally produced results of comparable accuracy. No statistically significant change was seen when meta-learning was introduced.

Classif	icatio	n - T-S	Stage)			914			are.		
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	49.40	76.67	73.70	72.40	67.27	76.27	75.47	78.27	76.90	74.57	75.97	74.30
В	51.67	76.67	75.20	73.40	71.43	74.90	76.67	76.67	76.67	74.57	75.97	74.30
С	69.57	76.67	76.67	76.67	69.93	73.93	74.13	77.40	76.67	78.23	73.50	74.70
D	69.57	76.67	76.67	76.67	69.93	73.93	74.13	77.40	76.67	78.23	73.50	74.70

Figure 156: T-Stage - Accuracy Results (Percentage)

Classif	ication	- T-St	age - /	AdaBo	ostM1						
Data Set	LGR	1R.AB1	J48.AB	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	49.40	74.77	67.10	67.27	59.30	51.57	78.27	75.53	74.87	74.47	75.07
В	51.67	74.13	66.50	71.43	62.13	63.47	76.67	71.97	71.03	73.90	74.37
С	69.57	75.40	70.80	69.93	65.83	65.73	75.83	75.90	76.10	74.53	75.27
D	69.57	75.40	70.80	69.93	65.83	65.73	75.83	75.90	76.10	74.53	75.27

Figure 157: T-Stage - Accuracy Results (Percentage) - AdaBoostM1

Classif	Classification - T-Stage - Bagging														
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG				
A	49.40	76.50	75.47	67.80	70.53	74.07	76.67	76.67	75.37	74.77	74.17				
В	51.67	76.67	76.27	71.40	73.77	75.67	76.67	76.67	75.37	74.77	73.97				
С	69.57	76.67	76.27	68.57	73.33	75.90	77.63	76.47	78.60	76.27	77.47				
D	69.57	76.67	76.27	68.57	73.33	75.90	77.63	76.47	78.60	76.27	77.47				

Figure 158: T-Stage - Accuracy Results (Percentage) - Bagging



Figure 159: T-Stage - Results Graph



Figure 160: T-Stage - Results Graph - AdaBoostM1



Figure 161: T-Stage - Results Graph - Bagging

Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	+	+	+	+	+	+	+	+	+	+	+
В	+	+	+	+	+	+	+	+	+	+	+
С	=	=	=));=:	=	H 3	=	=5	=	=3	=
D	Ŧ	=	=	=	=	Ŧ	=	=	=	=	=

Figure 162: T-Stage - T-Test vs. Logistic Regression

Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	+	+	+	=	=	+	+	+	+	+
В	+	+	+	=	=	+	+	+	+	+
С	=	=	=	=	=	=	H	=	=	=
D	=	=	=	=	=	=	=	=	=	=

Figure 163: T-Stage - T-Test vs. Logistic Regression - AdaBoostM1

Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	+	+	÷	+	+	+	+	+	+	+
В	+	+	+	+	+	+	+	+	+	+
С	=	=	=	=:	=	=	=	=	=	=
D	=	=	H	=	=	=	=	=	H	=

Figure 164: T-Stage - T-Test vs. Logistic Regression - Bagging

5.3 C3 - N-Stage

For the n-staging tests among N=51 patients, we predict n-stage of 3 classes which are dominated by value N1 (approx. 2:1 ratio to remaining values). Distribute of target values is shown in Table 8. Classification accuracy for n-size prediction generally ranges from 55% to 85%. The majority of algorithms in the original A data sets performed better than logistic regression via t-testing-particulary k-nearest-neighbor, locally-weighted-learning, and Bayesian nets. For the remaining data sets, algorithms generally performed equally. Data sets with supervised discretization and attribute selection generally produced results with higher accuracy. No statistically significant change was seen when meta-learning was introduced.

Classif	icatio	n - N-S	Stage									
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	54.97	66.80	62.10	54.10	70.53	70.57	67.87	67.57	66.60	63.73	68.60	64.87
В	58.60	66.80	67.93	55.27	69.90	66.23	66.43	66.47	64.70	63.73	68.60	64.87
С	73.83	66.70	70.00	60.67	70.80	74.30	72.73	78.57	80.10	82.87	82.30	83.30
D	73.83	66.70	70.00	60.67	70.80	74.30	72.73	78.57	80.10	82.87	82.30	83.30

Figure 165: N-Stage - Accuracy Results (Percentage)

Classif	fication	1 - N-S	tage - /	AdaBo	oostM1						
Data Set	LGR	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	54.97	62.27	57.43	70.53	67.50	62.07	67.57	62.83	63.10	66.70	60.13
В	58.60	66.07	55.77	69.90	62.07	65.70	66.47	64.70	63.47	67.90	59.33
С	73.83	70.30	71.03	70.80	72.00	77.47	73.03	79.50	81.17	78.73	79.17
D	73.83	70.30	71.03	70.80	72.00	77.47	73.03	79.50	81.17	78.73	79.17

Figure 166: N-Stage - Accuracy Results (Percentage) - AdaBoostM1

Classif	ication	- N-St	age -	Baggi	ng						
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	54.97	64.83	64.03	70.73	70.57	68.83	69.20	66.77	62.63	69.73	69.00
В	58.60	67.43	64.60	68.27	66.27	66.23	67.43	64.53	62.80	70.27	68.53
С	73.83	70.57	67.47	70.23	71.97	73.13	76.80	79.87	83.10	80.20	79.87
D	73.83	70.57	67.47	70.23	71.97	73.13	76.80	79.87	83.10	80.20	79.87

Figure 167: N-Stage - Accuracy Results (Percentage) - Bagging



Figure 168: N-Stage - Results Graph



Figure 169: N-Stage - Results Graph - AdaBoostM1



Figure 170: N-Stage - Results Graph - Bagging

Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	÷	=	=	+	+	÷	+	+	=	+	=
В	=	=	=	=	=	Ξ.	=	=	=	=	=
С	=	=	=	=	=	=	=	= 5	=	=	=
D	=	=	=	=	=	=	=	=	=	=	=

Figure 171: N-Stage - T-Test vs. Logistic Regression

Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	=	=	+	+	=	+	=	=	=	=
В	=	=	=	=	= 2	=	=	=	=	=:
С	=	=	=	=	= 3	=	=	=	=	=
D	=	Ξ	=	=	=	=	=	1 1	=	=

Figure 172: N-Stage - T-Test vs. Logistic Regression - AdaBoostM1

N-Stage										
Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BC	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	=	=	+	+	+	+	+	=	+	+
В	=	=	=	=	=	=	=	=	+	=
С	=	=	=	=	I	=	=	=	=	=
D	=	=	=	=	=	=	=	=	=	=
+ : Superio	or to LGR	= : Equ	uivalent	to LGR	- : Inferi	ior to LO	€R			·

Figure 173: N-Stage - T-Test vs. Logistic Regression - Bagging

5.4 C4 - Vascular Involvement

For the vascular involvement tests among N=74 patients, we predict the values of 2 classes which are dominated by 'false' values (approx. 80% of patients). Distribute of target values is shown in Table 9. Classification accuracy for vascular involvement prediction generally ranges from 75% to 85%. Analysis of the associated confusion matrices show that prediction dominates for the majority 'false' class and under-predicts the remaining values. Data sets with supervised discretization and attribute selection generally produced results with higher accuracy. No statistically significant change was seen when meta-learning was introduced.

Classif	icatio	n - Va	scula	ture								
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
Α	75.51	85.78	80.56	79.72	77.32	84.23	79.28	83.80	78.48	77.17	79.93	78.82
В	77.24	85.78	82.26	79.82	81.82	85.33	83.91	79.73	77.87	76.83	79.81	79.13
С	84.39	85.78	83.91	84.78	82.90	83.89	84.00	85.74	86.96	86.60	84.52	84.53
D	84.39	85.78	83.91	84.78	82.90	83.89	84.00	85.74	86.96	86.60	84.52	84.53

Figure 174: Vascular Involvement - Accuracy Results (Percentage)

Classif	ication -	Vascu	lature	- Ada	Boost	M1	410				
Data Set	LGR	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	75.51	86.33	84.56	77.32	74.67	74.00	82.37	82.97	82.46	80.40	80.38
В	77.24	81.27	80.12	81.82	74.94	73.72	79.98	78.18	77.83	76.94	79.56
С	84.39	85.19	78.94	80.69	78.48	80.38	82.77	85.39	85.93	83.00	82.12
D	84.39	85.19	78.94	80.69	78.48	80.38	82.77	85.39	85.93	83.00	82.12

Figure 175: Vascular Involvement - Accuracy Results (Percentage) - AdaBoostM1

Classif	ication -	Vascu	lature	- Bag	ging						
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	75.51	85.22	84.13	78.09	79.50	82.06	85.99	82.78	81.60	82.71	81.14
В	77.24	85.23	84.12	81.73	81.93	82.82	81.38	79.29	78.94	80.92	79.68
С	84.39	85.22	84.57	83.44	84.32	84.33	85.19	87.39	86.82	85.86	85.63
D	84.39	85.22	84.57	83.44	84.32	84.33	85.19	87.39	86.82	85.86	85.63

Figure 176: Vascular Involvement - Accuracy Results (Percentage) - Bagging



Figure 177: Vascular Involvement - Results Graph



Figure 178: Vascular Involvement - Results Graph - AdaBoostM1



Figure 179: Vascular Involvement - Results Graph - Bagging

Vascul	Vasculature T-Test - ML Algorithms vs. Logistic Regression														
Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3				
A	+	=	=	=	=	=	=	=	=	=	=				
В	+	=	=	=	=	=	=	=	=	=	=				
С	=	=	=	=	=	=	=	=	=	=	=				
D	=	=	=	=	=	=	=	=	=	=	=				
+ : Superi	or to LG	+ : Superior to LGR = : Equivalent to LGR - : Inferior to LGR													

Figure 180: Vascular Involvement - T-Test vs. Logistic Regression

Vascul	ature T	-Test - I	ML Alg	s. w/	AdaBo	ostM1	vs. Log	istic Re	gressio	on
Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	+	=	=	=	=	=	=	=	=	=
В	=	=	=	=	2 E	=	=	E	=	=
С	=	=	=	=	=	=	=	=	=	=
D	=	=	=	=	Ξ	=	=	=	=	=
+ : Superi	or to LGR	= : Equiv	alent to l	LGR -:	Inferior t	o LGR	ļ	2	<u>6</u>	2

Figure 181: Vascular Involvement - T-Test vs. Logistic Regression - AdaBoostM1

Vascul	Vasculature T-Test - ML Algs. w/ Bagging vs. Logistic Regression													
Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BC	NVB.BG	BN1.BG	BN2.BG	BN3.BG				
A		=	=	=	=	=	= 2	=	Ξ.	Ξ				
В	=	=	=	=	=	=	=)	=	E	÷				
С	=	=	=	=8	=	=	=:	=	=	=				
D	=	=	=	=	=	=	=	=	Ŧ	E				

Figure 182: Vascular Involvement - T-Test vs. Logistic Regression - Bagging

5.5 C5 - Histology

For the histology tests among N=91 patients, we predict value of 14 target class values which are dominated by 'Adenocarcinoma of Pancreas - NOS' and 'Ductal Adenocarcinoma of Pancreas' (these histology values dominate approximately 55% of instances). Distribute of target values is shown in Table 10. Classification accuracy for histology prediction models generally range from 35% to 55%. Analysis of the associated confusion matrices show that prediction dominates for the majority classes and 'IPMN - Benign or CiS' while underpredicting the remaining values. Data sets with supervised discretization and attribute selection combined with Bayesian net predictions generally produced results with higher accuracy than logistic regression via t-tests. Remaining machine learning algorithms were comparable to logistic regression accuracy in most cases. No statistically significant change was seen when meta-learning was introduced. High-performance models based on histology classification are presented in Section 6.1.

Classif	icatio	n - His	stolog	IV								
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	45.69	29.73	35.12	39.41	37.71	36.69	40.52	45.34	42.91	51.52	52.89	51.18
В	43.79	29.73	35.78	40.07	43.81	41.67	48.59	50.00	51.52	51.41	52.89	51.07
С	41.59	29.73	35.78	42.93	47.57	47.91	49.96	53.52	56.07	54.39	50.36	49.46
D	41.59	29.73	35.78	42.93	47.57	47.91	49.96	53.52	56.07	54.39	50.36	49.46

Classif	ication	- Hist	ology ·	Ada	BoostM	1					
Data Set	LGR	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	45.69	35.68	43.41	37.71	41.29	40.66	45.34	40.92	52.22	54.30	53.36
В	43.79	38.07	45.44	43.81	44.13	48.70	50.56	48.56	51.32	54.21	53.01
С	41.59	37.73	48.32	47.57	43.44	49.84	51.82	48.74	50.19	47.13	47.59
D	41.59	37.73	48.32	47.57	43.44	49.84	51.82	48.74	50.19	47.13	47.59

Figure 184: Histology - Accuracy Results (Percentage) - AdaBoostM1

Classif	ication	- His	tology	/ - Bag	ging						
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	45.69	43.89	46.87	36.49	38.77	43.44	44.36	41.39	51.66	51.54	50.69
В	43.79	43.54	47.07	42.37	43.81	44.88	51.89	52.77	51.98	52.77	51.21
С	41.59	43.43	48.83	46.13	49.77	52.18	53.77	54.57	53.40	51.23	50.48
D	41.59	43.43	48.83	46.13	49.77	52.18	53.77	54.57	53.40	51.23	50.48

Figure 185: Histology - Accuracy Results (Percentage) - Bagging



Figure 186: Histology - Results Graph



Figure 187: Histology - Results Graph - AdaBoostM1



Figure 188: Histology - Results Graph - Bagging

Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	-	=	=	=	=		=	=) (1) (1)	#	=
В	 .	=	=	=	=	=	=	=	=	=	=
С		=	=	=	=	=	=	+	+	=	=
D	-	=	=	=	=	=	=	+	+	=	=

Figure 189: Histology - T-Test vs. Logistic Regression

Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	=	=	=	=	=	=	=	=	=	=
В	=	=	÷)	=	=	=	=	=	=	=
С	=	=	≘	=	=	=	=	=	=	=
D	=	=	Ξo	=	=	=	=	=	=	

Figure 190: Histology - T-Test vs. Logistic Regression - AdaBoostM1

Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BC	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	=	=	=3	=	=	= 0	=	=	=	÷.
В	=	=	=	=	=	Ξ	=	=	=	=
С	=	=	=	=	=	+	+	=:	=	=
D	=	-	=	=	=	+	+	=	=	=

Figure 191: Histology - T-Test vs. Logistic Regression - Bagging

5.6 C6 - Malignancy

For the malignancy tests among N=91 patients, we predict value of 2 classes which are dominated by 'Malignant' values (approx. 80% of cases). Distribute of target values is shown in Table 11. Classification accuracy for malignancy prediction generally ranges from 70% to 85%. Analysis of the associated confusion matrices show a reasonable spread between the majority classes and minority values. Data sets with supervised discretization and attribute selection generally produced results with higher accuracy. Classification algorithms were t-test comparable to logistic regression accuracy in most cases. No statistically significant change was seen when meta-learning was introduced.

Classif	icatio	n - Ma	ligna	ncy								
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	76.90	82.44	84.64	84.44	70.56	79.14	77.38	64.39	57.97	75.42	74.10	74.63
В	75.38	82.44	79.68	79.07	72.51	78.82	77.40	80.31	76.18	72.12	71.56	72.84
С	83.18	82.44	80.44	82.44	82.87	84.61	84.61	85.63	81.56	83.54	83.84	83.51
D	83.18	82.44	80.44	82.44	82.87	84.61	84.61	85.63	81.56	83.54	83.84	83.51

Figure 192: Malignancy - Accuracy Results (Percentage)

Classif	ication -	Malig	nancy	- Ada	BoostN	/11					
Data Set	LGR	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	76.90	76.48	80.54	70.56	63.38	75.93	72.36	69.38	81.13	78.12	75.93
В	75.38	75.20	74.52	72.51	73.62	69.54	77.80	76.70	77.93	74.36	73.73
С	83.18	81.57	80.30	80.33	77.79	81.68	82.74	84.84	85.08	83.87	83.22
D	83.18	81.57	80.30	80.33	77.79	81.68	82.74	84.84	85.08	83.87	83.22

Figure 193: Malignancy - Accuracy Results (Percentage) - AdaBoostM1

Classif	Classification - Malignancy - Bagging													
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG			
A	76.90	82.33	85.54	70.54	74.50	76.17	70.41	66.88	75.66	77.73	76.44			
В	75.38	81.22	80.27	71.40	74.41	76.17	79.64	76.63	72.49	75.96	74.79			
С	83.18	81.67	81.57	82.53	83.62	84.83	84.74	81.46	82.88	85.94	85.07			
D	83.18	81.67	81.57	82.53	83.62	84.83	84.74	81.46	82.88	85.94	85.07			

Figure 194: Malignancy - Accuracy Results (Percentage) - Bagging



Figure 195: Malignancy - Results Graph



Figure 196: Malignancy - Results Graph - AdaBoostM1



Figure 197: Malignancy - Results Graph - Bagging

Maligna	Malignancy T-Test - ML Algorithms vs. Logistic Regression												
Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3		
A	=	=	=	=	=	=	-	-	=	=	=		
В	=	=	=	=	=	=	=	=	=	=	=		
С	=	=	=	=	=	=	=	=	=	=	=		
D	=	=	=	=	=	=	=	=	=	=	=		
+ : Superio	or to LG	R =:E	quivale	nt to LG	GR -∶li	nferior t	o LGR						

Figure 198: Malignancy - T-Test vs. Logistic Regression

Malign	Malignancy T-Test - ML Algs. w/ AdaBoostM1 vs. Logistic Regression											
Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1		
A	=	=	=	=	=	=	=	=	=	÷.		
В	=	=	=	=	S. H.	=	=	=	=	=		
С	=	=	=	=	=	=	=	=	=	=		
D	=	=	=	=	E	=	=	=	=	=		
+ : Superi	ior to LGR	= : Equiv	alent to	LGR -:	Inferior t	o LGR	1	8	1999 - 19			

Figure 199: Malignancy - T-Test vs. Logistic Regression - AdaBoostM1

Malignancy T-Test - ML Algs. w/ Bagging vs. Logistic Regression											
Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG	
A	=	+	= 8	=	=	=	=	=	=	=	
В	=	=	=	=	1	=	2 2	=	=	=	
С	=	=	=	=	=	=	=	=	=	=	
D	=	=	=	=	Ξ	=	÷	=	=	=	
+ : Superi	or to LGR	t = : Equ	ivalent	to LGR ·	: Inferio	to LGR	3	38	6		

Figure 200: Malignancy - T-Test vs. Logistic Regression - Bagging

5.7 C7 - ECOG 6-Month

For ECOG 6-Month tests among N=72 patients, we predict value of 3 classes which are reasonably well-distributed (ECOG values represented are those available among instances.). Distribution of target values is shown in Table 12. Classification accuracy for ECOG prediction generally ranges from 55% to 75%. Data sets with supervised discretization and attribute selection generally produced results with higher accuracy. Classification algorithms were t-test comparable to logistic regression accuracy in most cases. No statistically significant change was seen when meta-learning was introduced.

Classif	Classification - ECOG 6-Month												
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3	
A	49.20	51.43	48.21	49.68	55.16	57.34	54.95	51.55	54.09	63.54	55.43	54.30	
В	53.21	51.43	54.84	52.34	53.12	56.71	56.86	66.16	65.23	63.37	55.43	54.73	
С	64.93	51.43	55.37	59.43	68.27	69.80	68.80	72.29	69.64	70.89	70.30	68.21	
D	64.93	51.43	55.37	59.43	68.27	69.80	68.80	72.29	69.64	70.89	70.30	68.21	

Figure 201:	ECOG 6-Month -	Accuracy Results	(Percentage)
-------------	----------------	------------------	--------------

Classif	Classification - ECOG 6-Month - AdaBoostM1												
Data Set	LGR	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1		
A	49.20	53.89	50.82	55.16	54.68	55.75	50.59	45.82	58.21	50.34	56.86		
В	53.21	55.20	49.77	53.12	55.96	55.09	57.32	54.96	58.82	51.00	51.77		
С	64.93	57.91	57.86	59.84	70.07	72.21	62.32	66.25	66.29	64.79	61.05		
D	64.93	57.91	57.86	59.84	70.07	72.21	62.32	66.25	66.29	64.79	61.05		

Figure 202: ECOG 6-Month - Accuracy Results (Percentage) - AdaBoostM1

Classif	Classification - ECOG 6-Month - Bagging													
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG			
A	49.20	53.18	54.00	55.37	54.86	57.02	53.14	55.21	63.12	58.14	54.05			
В	53.21	56.70	55.93	55.89	57.80	57.95	66.57	66.23	63.77	58.71	55.36			
С	64.93	58.75	61.45	68.79	69.23	69.68	72.14	69.96	70.21	69.43	69.91			
D	64.93	58.75	61.45	68.79	69.23	69.68	72.14	69.96	70.21	69.43	69.91			

Figure 203: ECOG 6-Month - Accuracy Results (Percentage) - Bagging



Figure 204: ECOG 6-Month - Results Graph



Figure 205: ECOG 6-Month - Results Graph - AdaBoostM1


Figure 206: ECOG 6-Month - Results Graph - Bagging

ECOG	ECOG 6-Month T-Test - ML Algorithms vs. Logistic Regression													
Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3			
A	=	=	=	=	=	=	=	=	=	=	=			
В	=	=	=	=	=	=	=	=	=	=	=			
С	=	=	=	=	=	=	=	=	Ξ.	=	=			
D	=	=	=	=	=	=	=	=	=	=	=			
+ : Superi	or to LG	+ : Superior to LGR = : Equivalent to LGR - : Inferior to LGR												

Figure 207: ECOG 6-Month - T-Test vs. Logistic Regression

ECOG	6-Month	n T-Tes	t - ML	Algs.	w/ Ada	Boost	M1 vs. L	ogistic	Regres	ssion
Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	=	=	=	=	=	=	=:	=	*	=:
В	=	=	=	=	=	=	=	=	=	=
С	=	=	=	=	= :	=	π.	=	=	=
D	=	=	=	=	=	=	=×	=	=	=
+ : Superi	ior to LGR	= : Equiv	alent to	LGR -:	Inferior t	o LGR	8	10	2	61

Figure 208: ECOG 6-Month - T-Test vs. Logistic Regression - AdaBoostM1

Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	BN1.BG	BN2.BG	BN3.BG	NVB.BG
A	=	=	=	=:	=	=	z :	Ξ.	=	=
В	=	=	=	=	=	=	=	=	=	=
С	=	=	=	(E)	=	=	=	=	=	=
D	=	=	=	=	=	=	=	=	=	=

Figure 209: ECOG 6-Month - T-Test vs. Logistic Regression - Bagging

5.8 C8 - ECOG 9-Month

For ECOG 9-Month tests among N=57 patients, we predict value of 4 classes which are reasonably well-distributed (ECOG values represented are those available among instances.). Distribute of target values is shown in Table 13. Classification accuracy for ECOG prediction generally ranges from 45% to 70%. Data sets with supervised discretization and attribute selection generally produced results with higher accuracy. Classification algorithms were t-test comparable to logistic regression accuracy in most cases. No statistically significant change was seen when meta-learning was introduced.

Classifi	cation	ı - EC	OG 9	-Mont	th							
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	45.57	58.10	46.90	47.47	44.07	56.33	52.43	59.53	60.07	49.67	53.57	51.87
В	42.63	58.10	49.80	50.57	43.47	53.70	51.37	54.97	54.53	49.50	53.57	51.87
С	62.20	58.10	50.97	52.67	58.80	54.47	55.53	64.80	70.33	70.20	63.87	64.00
D	62.20	58.10	50.97	52.67	58.80	54.47	55.53	64.80	70.33	70.20	63.87	64.00

Figure 210: ECOG 9-Month - Accuracy Results (Percentage)

Classif	ication	- ECO	G 9-N	Ionth -	AdaBo	ostM1					
Data Set	LGR	1R.AB1	J48.AB	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	45.57	48.20	42.80	44.07	37.60	49.07	59.53	59.10	50.30	45.67	47.57
В	42.63	51.77	43.00	43.47	32.60	51.17	52.70	45.53	49.17	44.13	48.13
С	62.20	51.90	50.40	58.80	52.43	58.23	53.10	57.03	53.87	59.00	57.90
D	62.20	51.90	50.40	58.80	52.43	58.23	53.10	57.03	53.87	59.00	57.90

Figure 211: ECOG 9-Month - Accuracy Results (Percentage) - AdaBoostM1

Classif	ication	- ECO	G 9-M	onth -	Bagg	jing					
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	45.57	55.20	53.80	42.57	45.57	50.03	60.80	61.17	49.73	51.77	49.13
В	42.63	53.87	54.60	38.53	44.13	51.30	55.37	54.53	49.70	52.47	49.27
С	62.20	54.67	57.20	54.27	57.17	55.70	64.70	70.03	70.63	62.53	61.40
D	62.20	54.67	57.20	54.27	57.17	55.70	64.70	70.03	70.63	62.53	61.40

Figure 212: ECOG 9-Month - Accuracy Results (Percentage) - Bagging



Figure 213: ECOG 9-Month - Results Graph



Figure 214: ECOG 9-Month - Results Graph - AdaBoostM1



Figure 215: ECOG 9-Month - Results Graph - Bagging

Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	E	=	=	E	=	=	+	+	=	=	=
В	+	=	=	=	=	=	=	=	=	=	=
С	=	=	=	=	=	=	=	=	=	=	=
D	=	=	=	=	=	=	=	=	=	=	=

Figure 216: ECOG 9-Month - T-Test vs. Logistic Regression

ECOG 9-Month T-Test - ML Algs. w/ AdaBoostM1 vs. Logistic Regression													
Data Set	1R.AB1	J48.AB1	IB1.AB	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1			
A	Þ	=	=	=	=	+	+	=	=	÷			
В	=	3	=	E.	=	=	E	E	=	ŧ			
С	=	=	=	.≓¢	=	=::	=	=	=	=			
D	÷	=	Ξ	=	=	=	=	2	=	=			

Figure 217: ECOG 9-Month - T-Test vs. Logistic Regression - AdaBoostM1

Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	=	=	=	=	=	+	+	=	= 8	= 8
В	=	2	=	=	=	=		=	=	=
С	=	=	=	=	=	=	=	=	=0	=
D	=	=	=	=	=	=	=	=	=	=

Figure 218: ECOG 9-Month - T-Test vs. Logistic Regression - Bagging

5.9 C9 - ECOG 12-Month

For ECOG 12-Month tests among N=44 patients, we predict value of 4 classes which are reasonable well distributed (ECOG values represented are those available among instances.). Distribute of target values is shown in Table 14. Classification accuracy for ECOG prediction generally ranges from 35% to 55%. The majority of algorithms in A and B data sets performed better than logistic regression via t-testing; again, this seems due more to logistic regression's poor performance on these sets. Data sets with supervised discretization and attribute selection generally produced results with equivalent accuracy. Classification algorithms in C and D sets were t-test comparable to logistic regression accuracy in most cases. No statistically significant change was seen when meta-learning was introduced.

Classif	icatio	n - EC	OG 1	2-Moi	nth							
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	37.90	52.20	47.05	41.25	42.45	48.45	48.35	48.15	41.35	46.85	39.50	38.70
В	40.70	52.20	48.00	45.05	42.40	49.55	43.05	48.70	50.60	47.25	39.50	38.90
С	43.60	52.20	39.40	54.45	48.15	45.50	46.15	50.55	53.95	51.80	49.30	48.70
D	43.60	52.20	39.40	54.45	48.15	45.50	46.15	50.55	53.95	51.80	49.30	48.70

Figure 219: ECOG 12-Month - Accuracy Results (Percentage)

Classif	ication	- ECO	G 12-	Month	- AdaB	oostM1		<i>b</i>			<i></i>
Data Set	LGR	1R.AB1	J48.AB	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	37.90	43.25	41.95	40.70	42.80	43.25	46.00	43.30	46.25	42.50	42.45
В	40.70	43.50	38.55	38.70	35.95	38.60	46.95	47.60	45.25	42.40	40.00
С	43.60	41.10	51.10	43.80	45.00	45.75	49.55	51.05	50.90	47.55	46.90
D	43.60	41.10	51.10	43.80	45.00	45.75	49.55	51.05	50.90	47.55	46.90

Figure 220: ECOG 12-Month - Accuracy Results (Percentage) - AdaBoostM1

Classif	ication	- EC	OG 12-	Mont	h - Ba	gging					
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	37.90	48.40	45.10	43.20	43.35	45.50	44.50	39.85	46.05	39.45	38.00
В	40.70	47.65	47.40	38.80	40.30	43.80	50.25	49.70	46.55	40.95	38.15
С	43.60	43.25	57.00	48.45	46.90	44.30	50.20	55.45	51.70	46.60	47.40
D	43.60	43.25	57.00	48.45	46.90	44.30	50.20	55.45	51.70	46.60	47.40

Figure 221: ECOG 12-Month - Accuracy Results (Percentage) - Bagging



Figure 222: ECOG 12-Month - Results Graph



Figure 223: ECOG 12-Month - Results Graph - AdaBoostM1



Figure 224: ECOG 12-Month - Results Graph - Bagging

Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	=	=	=	=	=	=	=		=	=	=
B	=	=	=	=	=	=	=	=	=	=	=
С	=	=	=	=	=	=	=	=	=	=	=
D	=	=	=	=	=	=	=	=	=	=	=

Figure 225: ECOG 12-Month - T-Test vs. Logistic Regression

ECOG	12-Mo	nth T-T	'est - l	ML Alg	s. w/ A	daBoos	tM1 vs	Logist	ic Regr	ession
Data Set	1R.AB1	J48.AB1	IB1.AB	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	=	=	=	=	=	=	=	=	=	=
В	=	E.	=	=	=	=	=		E	=
С	=	=	=	=	=	=	=	=	=	=
D	=	2	=	=	=	=	=	8	=	=
+ : Superi	ior to LGI	R = : Equ	ivalent	to LGR	: Inferior	to LGR	200	5°		

Figure 226: ECOG 12-Month - T-Test vs. Logistic Regression - AdaBoostM1

ECOG	12-Mo	nth T-	Test -	ML AI	gs. w	/ Baggi	ing vs.	Logis	tic Reg	gression
Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	=	=	# 3	=	=	Ξ	=	=	=	=
В	=	=	=	=	=	=	=	E	E	=
С	=	=	=8	=	=	=	=	=	=	=
D	=	=	=	=	=	=	=	=	=	Ŧ

Figure 227: ECOG 12-Month - T-Test vs. Logistic Regression - Bagging

5.10 C10 - Survival

For survival tests among N=60 patients, we predict value of 4 numeric ranges which are evenly distributed between bins. Distribute of target values is shown in Table 15. Classification accuracy for survival prediction generally ranges from 40% to 60%. Naive Bayes and Bayesian nets in A and B data sets performed better than logistic regression via t-testing–a notable result. Data sets with supervised discretization and attribute selection generally produced results with higher accuracy. Remaining machine learning algorithms were t-test comparable to logistic regression accuracy in most cases. No statistically significant change was seen when meta-learning was introduced.

Classif	icatio	n - Su	rvival									
Data Set	LGR	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	41.67	33.33	26.17	46.00	50.33	46.17	33.67	45.83	40.83	43.33	47.50	47.50
В	39.33	33.33	31.83	45.00	47.67	39.50	34.50	47.00	41.17	43.33	48.50	48.00
С	42.50	33.33	34.17	43.50	55.67	54.83	56.00	52.67	57.17	56.67	55.67	52.00
D	42.50	33.33	34.17	43.50	55.67	54.83	56.00	52.67	57.17	56.67	55.67	52.00

Figure 228: Survival - Accuracy Results (Percentage)

Classif	ication	- Surv	vival -	AdaBo	ostM1						
Data Set	LGR	1R.AB1	J48.AB	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	41.67	32.33	43.17	50.33	41.33	33.67	45.83	42.50	45.00	46.83	47.33
В	39.33	32.83	42.00	47.67	39.33	33.83	47.00	38.83	43.17	47.83	48.50
С	42.50	34.00	44.50	54.33	52.50	55.50	53.17	54.67	53.83	53.17	49.00
D	42.50	34.00	44.50	54.33	52.50	55.50	53.17	54.67	53.83	53.17	49.00

Figure 229: Survival - Accuracy Results (Percentage) - AdaBoostM1

Classif	ication	- Surv	ival - I	Baggi	ng						
Data Set	LGR	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	41.67	34.33	43.50	47.83	40.50	38.17	42.17	40.17	42.00	46.33	46.33
В	39.33	34.17	44.67	44.50	43.17	40.67	46.33	40.00	40.50	47.50	46.83
С	42.50	35.50	45.67	51.50	54.00	53.50	52.83	55.83	55.00	53.00	53.00
D	42.50	35.50	45.67	51.50	54.00	53.50	52.83	55.83	55.00	53.00	53.00

Figure 230: Survival - Accuracy Results (Percentage) - Bagging



Figure 231: Survival - Results Graph



Figure 232: Survival - Results Graph - AdaBoostM1



Figure 233: Survival - Results Graph - Bagging

Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	E:	2	=	=	—) =) = a)) =s	=	=)=:
В	=	=	=	=	=	=	=	=	=	=	=
С	=	=	=	÷.	=	=	=	+	=	= :	=
D	=	=	=	=	=) == :	=	+	=	=	=

Figure 234: Survival - T-Test vs. Logistic Regression

	ssion	c Regre	Logisti	tM1 vs.	aBoos	w/ Ad	Algs.	st - ML	al T-Te	Surviva
1 BN3.AB1	BN2.AB1	BN1.AB1	NVB.AB1	LWL.AB1	IB3.AB1	IB2.AB1	IB1.AB1	J48.AB1	1R.AB1	Data Set
=	=	=	Ξ	=	=	±.	±`	=	=	A
=	=	=	÷.	=	=	=	2)	=	=	В
=	=	=	=	± l	=	=	=	=	=	С
=	=	=	=	=	=	=	=	=	=	D
-	=	E	H	= to LGR	= : Inferior	= to LGR ·	= ivalent 1	= = : Equ	= or to LGR	D +:Superi

Figure 235: Survival - T-Test vs. Logistic Regression - AdaBoostM1

Surviva	al T-Tes	st - ML	Algs.	w/ Ba	gging	yvs. Lo	gistic	Regre	ssion	
Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG
A	=	=	÷.	=	=	=	=	=	=	=
В		=	Ξ.	=	=	=	=	=	=	=
С	=	=	Ξ	=	=	=	=	=	=	=
D	=	=	24	=	=	=	=	=	=	=

Figure 236: Survival - T-Test vs. Logistic Regression - Bagging

5.11 R1 - Tumor Size

For tumor-size regression tests among N=74 patients, we predict numeric values ranging from 0 to 11 cm. Distribute of target values is shown in Table 6. Regression r-squared values for survival prediction range from .00 to .45. Linear regression and M5 model trees performed best. Data sets with attribute selection generally produced results with comparable r-squared values. Remaining machine learning algorithms were t-test inferior to linear regression accuracy in most cases. Meta-learning introduced statistically superior performance in multi-layer perceptrons when compared to linear regression performance.

Regres	sion - T	umor :	Size				
Data Set	LNR	ZR	M5M	M5R	RBF	MLP2	MLP3
E	0.27	0.00	0.28	0.29	0.03	0.17	0.17
F	0.41	0.00	0.37	0.28	0.07	0.32	0.32

Figure 237: Tumor Size - R-Squared Results

Regres	sion - 1	umor Si	ze - Ba	gging a	nd Stack	king			
Data Set	LNR	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK	
E	0.2	7 0.29	0.36	0.28	0.03	0.26	0.24		0.15
F	0.4	1 0.45	0.30	0.42	0.17	0.41	0.38		0.20

Figure 238: Tumor Size - R-Squared Results - AdaBoostM1



Figure 239: Tumor Size - Regression Results Graph



Figure 240: Tumor Size - Regression Results Graph - Bagging and Stacking

		T-Test - vs. Line		ression	1	
Data Set	ZR	M5M	M5R	RBF	MLP2	MLP3
E	-	=	=	1	+	-
F	-	=	1	4	<u>i</u>	-
+ : Super	ior to Ll	NR =: Equ	uivalent to	LNR -:	Inferior to	LNR

Figure 241: Tumor Size - T-Test vs. Linear Regression

	Size T- a-Learn			rithms Regress	ion		
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK
E	=	=	=	5	=	=	-
F	=	-	æ	-	=	=	-
+ : Superi	ior to LNR	= : Equiv	alent to LN	IR -:Infer	ior to LNR	l:	

Figure 242: Tumor Size - T-Test vs. Linear Regression - Meta-learners

5.12 R2 - ECOG 6-Month

For ECOG 6-Month regression tests among N=72 patients, we predict numeric values ranging from 0 to 2 (ECOG values represented are those available among instances.). Distribute of target values is shown in Table 12. Regression r-squared values for ECOG prediction range from .00 to .27. Multi-layer perceptrons and RFB networks perform best, particularly with meta-learning on set F. Data sets with attribute selection generally produced results with higher r-squared values. Remaining machine learning algorithms were t-test inferior to linear regression accuracy in most cases. Meta-learning introduced statistically superior performance in multi-layer perceptrons, M5 model trees, and linear regression with bagging when compared to standard linear regression performance. High-performance models based on ECOG 6-Month regression are presented in Section 6.3.

Regression - ECOG 6-Month									
Data Set	LNR	ZR	M5M	M5R	RBF	MLP2	MLP3		
E	0.03	0.00	0.16	0.01	0.06	0.07	0.09		
F	0.26	0.00	0.23	0.23	0.14	0.17	0.27		

Figure 243: ECOG 6-Month - R-Squared Results

Regres	Regression - ECOG 6-Month - Bagging and Stacking								
Data Set	LNR	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK	
E	0.03	0.13	0.12	0.04	0.08	0.12	0.11	0.03	
F	0.26	0.25	0.23	0.32	0.24	0.31	0.32	0.18	

Figure 244: ECOG 6-Month - R-Squared Results - AdaBoostM1



Figure 245: ECOG 6-Month - Regression Results Graph



Figure 246: ECOG 6-Month - Regression Results Graph - Bagging and Stacking

				i.	
ZR	M5M	M5R	RBF	MLP2	MLP3
-	+	=	=	+	+
-) ±:	=	-	<u>5</u>	=
	hms v	hms vs. Line	hms vs. Linear Reg		hms vs. Linear Regression

Figure 247: ECOG 6-Month - T-Test vs. Linear Regression

ECOG	6-Month	T-Test	t - ML AI	gorithm	s				
w/ Meta	a-Learne	ers vs.	Linear R	egressi	on				
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK		
E	+	+	+	=	+	+	=		
F	=	=	+	=	=	+	-		
+ : Superie	+ : Superior to LNR = : Equivalent to LNR - : Inferior to LNR								

Figure 248: ECOG 6-Month - T-Test vs. Linear Regression - Meta-learners

5.13 R3 - ECOG 9-Month

For ECOG 9-Month regression tests among N=57 patients, we predict numeric values ranging from 0 to 3 (ECOG values represented are those available among instances.). Distribute of target values is shown in Table 13. Regression r-squared values for ECOG prediction range from .00 to .25. Multi-layer perceptrons and RFB networks perform best, particularly on set E. Data sets with attribute selection generally produced results with higher r-squared values. Remaining machine learning algorithms were t-test comparable or inferior to linear regression accuracy in most cases. Meta-learning introduced statistically superior performance in most tested models when compared to standard linear regression performance. High-performance models based on ECOG 9-Month regression are presented in Section 6.4.

Regression - ECOG 9-Month								
Data Set	LNR	ZR	M5M	M5R	RBF	MLP2	MLP3	
E	0.00	0.00	0.07	0.03	0.00	0.00	0.00	
F	0.04	0.00	0.10	0.00	0.12	0.08	0.10	

Figure 249: ECOG 9-Month - R-Squared Results

Regres	Regression - ECOG 9-Month - Bagging and Stacking								
Data Set	LNR	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK	
E	0.00	0.07	0.08	0.00	0.03	0.01	0.01	0.0	
F	0.04	0.13	0.14	0.08	0.25	0.16	0.13	0.0	

Figure 250: ECOG 9-Month - R-Squared Results - AdaBoostM1



Figure 251: ECOG 9-Month - Regression Results Graph



Figure 252: ECOG 9-Month - Regression Results Graph - Bagging and Stacking

	ECOG 9-Month T-Test - ML Algorithms vs. Linear Regression									
Data Set		-	M5R	RBF	MLP2	MLP3				
E	=	+	+	=	=	=				
F	-	=	-	=	=	=				
+ : Superio	or to LNR	= : Equiv	valent to L	NR -:Inf	ferior to L	NR				

Figure 253: ECOG 9-Month - T-Test vs. Linear Regression

ECOG 9	9-Month	T-Test	- ML AI	gorithm	s				
w/ Meta	a-Learne	ers vs.	Linear R	egressi	on				
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK		
E	+	+	+	+	+	+	=		
F	+	+	=	+	+	+	=		
+ : Superio	+ : Superior to LNR = : Equivalent to LNR - : Inferior to LNR								

Figure 254: ECOG 9-Month - T-Test vs. Linear Regression - Meta-learners

5.14 R4 - ECOG 12-Month

For ECOG 12-Month regression tests among N=44 patients, we predict numeric values ranging from 0 to 3 (ECOG values represented are those available among instances.). Distribute of target values is shown in Table 14. Regression r-squared values for ECOG prediction range from .00 to .28. Data sets with attribute selection generally produced results with higher r-squared values. Remaining machine learning algorithms were t-test comparable or inferior to linear regression accuracy in most cases. Meta-learning introduced statistically superior performance in multi-layer perceptrons on the data set E when compared to standard linear regression performance.

Regres	Regression - ECOG 12-Month								
Data Set	LNR	ZR	M5M	M5R	RBF	MLP2	MLP3		
E	0.00	0.00	0.03	0.00	0.01	0.07	0.05		
F	0.22	0.00	0.14	0.00	0.10	0.25	0.20		

Figure 255: ECOG 12-Month - R-Squared Results

Regres						Stacking			
Data Set	LNR	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK	
E	0.00	0.01	0.00	0.01	0.01	0.07	0.05		0.00
F	0.22	0.23	0.05	0.28	0.21	0.27	0.24		0.08

Figure 256: ECOG 12-Month - R-Squared Results - AdaBoostM1



Figure 257: ECOG 12-Month - Regression Results Graph



Figure 258: ECOG 12-Month - Regression Results Graph - Bagging and Stacking

				i.	
ZR	M5M	M5R	RBF	MLP2	MLP3
1 51			-	+	+
-	1	-2	-	=	=
	hms	hms vs. Line	hms vs. Linear Reg		hms vs. Linear Regression

Figure 259: ECOG 12-Month - T-Test vs. Linear Regression

				Algorith Regress			
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK
E	-	E.	÷.	Ŧ	+	+	E
F	=	<u>Le</u>	+	=	=	=	2
+ : Super	ior to LNR	= : Equiv	alent to LN	IR -: Infer	ior to LNR		

Figure 260: ECOG 12-Month - T-Test vs. Linear Regression - Meta-learners

5.15 R5 - Survival

For survival regression tests among N=60 patients, we predict numeric values ranging from 0.9 to 29.3 months. Distributions of target values is shown in Table 15. Regression r-squared values for survival prediction range from .00 to .28. Multi-layer perceptrons and linear regression with bagging performed generally better than linear regression, particularly on set E. Data sets with attribute selection generally produced results with higher r-squared values for meta-learning tests. Remaining machine learning algorithms had varied t-test accuracies when compared to linear regression. Meta-learning introduced instances of statistically inferior performance on both data sets.

Regres	sion - S	Burviva	I				
Data Set	LNR	ZR	M5M	M5R	RBF	MLP2	MLP3
E	0.01	0.00	0.00	0.00	0.01	0.08	0.05
F	0.25	0.00	0.00	0.00	0.02	0.26	0.25

Figure 261: Survival - R-Squared Results

Regres	sion - S	urvival -	Baggi	ng and \$	Stacking	l			
Data Set	LNR	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK	
E	0.01	0.01	0.00	0.02	0.00	0.07	0.07	(0.00
F	0.25	0.03	0.00	0.27	0.06	0.28	0.27	0	0.04

Figure 262: Survival - R-Squared Results - AdaBoostM1



Figure 263: Survival - Regression Results Graph



Figure 264: Survival - Regression Results Graph - Bagging and Stacking

		est - ML vs. Line	-	ression	Ĩ.	
Data Set		M5M	M5R	RBF	MLP2	MLP3
E	1 51	=		=	+	+
F	-	1	-2	-	=	=

Figure 265: Survival - T-Test vs. Linear Regression
	al T-Tes a-Learn			ms Regress	ion		
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK
E	=	=	+	÷	+	+	Ŧ
F		2	=		=	=	2

Figure 266: Survival - T-Test vs. Linear Regression - Meta-learners

6 High-Performance Predictive Models

Several of the high-performance machine learning models are described in this section. Two models from the classification experiments and two from the regression experiments are demonstrated. Each of these models outperform traditional regression methods via statistical tests. Each model also exhibits interesting structural characteristics, both in their internal design and the feature-selected attribute sets used to generate them. Verbatim Weka output of these models follows each section.

6.1 Classification - Histology - Data Set C - Bayesian Net 2-Parent

Shown here is a Bayesian Net 2-Parent classifier with high predictive accuracy for majority target class values. This model is taken from the C5 experiments in Section 5.5. Histology prediction is difficult given the wide variety of categorical possibilities (14 types are represented here). Additionally, certain histology types are only rarely represented in the clinical setting (MEN-I, pseudopapillary tumors, renal mets). As accurate prediction across all types is difficult, we seek instead to demonstrate models which can predict some of the more frequently occurring histologic values, including adenocarcinomas, neuroendocrine tumors, and IMPNs.

A graphical representation of this Bayes Net model is demonstrated in Figures 267 and 268. Classification accuracy for this particular Bayes Net model is 50.55%. For the three most frequently occurring histologic types, 'Adenocarcinoma of Pancreas - NOS', 'Ductal Adenocarcinoma of Pancreas', and 'IPMN - Benign or CiS', the predictive accuracy of this model is 79.03%. The Confusion matrix illustrated in Figure 269 illustrates the model's predictive accuracy for different histologic values, with the three majority histologic values shown boxed.

Experimental iterations of this data set with other Bayesian methods show that the accuracy can be pushed even higher. Naive Bayes classification retains the highest experiment accuracy at 56.07% (ref. Figure 183), although the Bayesian Net shown here exhibits a much more interesting probability structure. Each node on the Bayesian Net reflects the joint probability distribution for its related attribute as determined by the attribute values of its parent nodes. These probability distributions are determined by the comparative frequencies of attribute values within the data sets. Examples of these distributions are shown in Figure 270.

Feature-selection generated a 24 attribute subset for data sets C and D in these experiments. The field names and their explanations are listed in Table 24. Generally, experimental accuracy was much higher for feature-selected data sets. As this entire subset consists of categorical attributes, supervised discretization induces no change to the result set. Therefore, no experimental variation exists on models generated from data set C or D.

Field	Description
PresumptiveDx	Presumptive Diagnosis
SxWtloss	Presentation - Weight Loss
SxJaun	Presentation - Jaundice
SxNau	Presentation - Nausea
SxFati	Presentation - Fatigue
SxPru	Presentation - Pruritis
SxOT	Presentation - Other
CxDiab	Comorbidities - Diabetes
CTNodeOmit	CT - Nodal Omission
EUSVascOmit	EUS - Vascular Omission
EUSPortal	EUS - Portal Vein Involvement
EUSNoNode	EUS - No Nodal Involvement
EUSStagingT	EUS - T Staging
EUSCyto	EUS - Cytology
TxLap	Treatment - Laparoscopy
TxRadia	Treatment - Radiotherapy
TxChemo	Treatment - Chemotherapy
TxChemoGem	Treatment - Chemotherapy - Gemcitabine
ResPxType	Resection - Procedure Type
ResTransfusion	Resection - Transfusion
ResPOCourse	Resection - Postoperative Course
ResPathN	Resection - Pathology N-Stage
SurOncName	Surgical Oncologist
RadOncName	Radiation Oncologist

Table 24: Histology Feature-Selected Attribute Subset



Figure 267: Classification - Histology - Data Set C - Bayesian Net 2-Parent



Figure 268: Classification - Histology - Data Set C - Bayesian Net 2-Parent (continued)

															cix ===
a	b	С	d	е	£	g	h	i	j	k	1	m	n		< classified as
12	2	9	1	0	0	0	0	0	0	0	0	0	0	Ť	a = Adenocarcinoma_of_Pancreas/NOS
4	4	1	0	0	0	0	0	0	0	0	0	0	0	J,	b = Ampullary_Adenocarcinoma
8	0	18	0	0	0	0	0	0	0	0	1	0	0	1	<pre>c = Ductal_Adeno_of_Pancreas</pre>
1	0	1	2	0	0	0	1	0	0	0	0	0	0	Ì,	d = Neuroendocrine_(Islet_Cell/Carcinoid)
0	0	0	1	0	0	0	0	0	0	0	0	0	0	I.	e = Von_Hippel-Lindau_Syndrome
0	1	1	0	0	0	0	0	0	0	0	0	0	0	I.	f = Duodenal_Adenocarcinoma
0	0	0	0	0	1	0	0	0	0	0	0	0	0	I.	g = Distal_Cholangiocarcinoma
0	0	0	1	0	0	0	0	l	0	0	1	0	0	I.	h = Renal_Mets
0	0	1	0	0	0	0	0	1	0	0	2	0	0	Ť	i = Cystadenoma
0	0	ം	0	ം	0	0	െ	0	ି ()	0	1	ാ	0	J,	j = MEN-I
0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	k = Pseudopapillary_Tumor
0	0	1	1	0	0	0	0	0	0	0	9	0	0	I.	1 = IPMN/IPMTBenign_or_CiS
0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	m = Mucinous_Cystic_Neoplasm
0	0	0	0	0	0	0	0	1	0	0	0	0	0	T.	n = Benign Cyst

Figure 269: Classification - Histology - Data Set C - Confusion Matrix

	Histology	Presumpt	iveDx	FALSE	TRUE						
Aden	ocarcinoma_of_Pancreas/NOS	Pancreatic	Tumor	0.92		71					
	ocarcinoma_of_Pancreas/NOS	Periampulla		0.62	5 0.3	75					
	ocarcinoma_of_Pancreas/NOS	Suspicious_Par	ex. A service wards and a service serv	0.	5	0.5					
	ocarcinoma of Pancreas/NOS	IPMT/IF	No. of Concession, Name of	0.	5	0.5					
ENVIOUS-NOTION	ocarcinoma_of_Pancreas/NOS	Suspicious_Bile_	Duct Stricture	0.7	5 0	.25					
	ocarcinoma_of_Pancreas/NOS	Othe	Contract of the local division of the local	0.	5	0.5					
A	Ampullary_Adenocarcinoma			0.7		25					10 9
A	Ampullary_Adenocarcinoma	🐐 Probability D	istribution Tabl	e For EU	5Staging	Т					×
A	Ampullary_Adenocarcinoma	ŀ	listology	ł	USPortal	T4	T3	T2	T1	ΤX	6
A	Ampullary_Adenocarcinoma	Adenocarcinon	na_of_Pancrea:	s/NOS	TRUE	0.368	0.368	0.158	0.053	0.053	-
A	Ampullary_Adenocarcinoma		na_of_Pancreas		FALSE	0.128	0.333	0.385	0.128	0.026	
A	Ampullary_Adenocarcinoma		Adenocarcinom		TRUE	0.429	0.143	0.143	0.143	0.143	
E	Ductal_Adeno_of_Pancreas		Adenocarcinon		FALSE	0.048	0.619	0.143	0.143	0.048	
	Ductal_Adeno_of_Pancreas		eno_of_Pancrea		TRUE	0.2	0.2	0.2	0.2	0.2	
E	Ductal_Adeno_of_Pancreas		eno_of_Pancrea		FALSE	0.085	0.119	0.729	0.017	0.051	
	👙 Probability Distribution Table	For CTNodeOmit				×	0.2	0.2	0.2	0.2	
				1			0.067	0.6	0.2	0.067	
	Histology	SxPru	TRUE		ALSE	_	0.2	0,2	0.2	0.2	
leur	Adenocarcinoma_of_Pancreas		0.3	1		0.7	0.143	0.429	0.143	0.143	
leur	Adenocarcinoma_of_Pancreas		0.045		0.9		0,2	0.2	0.2	0.2	
leur	Ampullary_Adenocarcinom		0.562	1	0.4	and the second se	0.333	0.333	0.111	0.111	
leur	Ampullary_Adenocarcinom		0.167	-	0.8		0,2	0.2	0.2	0.2	
	Ductal_Adeno_of_Pancrea		0.241	1	0.7	1.7.4.	0.143	0.429	0,143	0.143	
	Ductal_Adeno_of_Pancrea	Contraction of the second s	0.25			75	0.2	0.2	0.2	0.2	
	Neuroendocrine_(Islet_Cell/Car		0.1	1		0.9	0.091	0.455	0.273	0.091	
	Neuroendocrine_(Islet_Cell/Car		0.25	1		.75	0.2	0.2	0.2	0.2	100000
	Von_Hippel-Lindau_Syndror		0.25	4		0.5	0.077	0.692	0.077	0.077	- CONTRACT
	Von_Hippel-Lindau_Syndroi		0.167	+	0.8		0.2	0.2	0.2	0.2	100000
	Duodenal_Adenocarcinom	the second se	0.107		10.20	0.5	0.143	0.429	0,143	0,143	- CONTRACT
	Duodenal_Adenocarcinom	Participant and a second	0.5	1		0.5	0.2	0.2	0.2	0.2	100000
	Distal_Cholangiocarcinom		0.25			.75	0.143	0.429	0,143	0,143	ЬM.
	Distal_Cholangiocarcinom	FALSE	0.125	1	0.8	A					
	Renal_Mets Renal Mets	TRUE	0.123		10.20	0.5					
	Cystadenoma	FALSE	0.7	1).3					
	Cystadenoma	TRUE	0.5	<u></u>		0.5					
	MEN-I	FALSE	0.25			.75					
	MENLI	TRUE	0.5			0.5					

Figure 270: Classification - Histology - Data Set C - Joint Probability Distribution Examples

Weka Output:

=== Run information ===

- Scheme: weka.classifiers.bayes.BayesNet -D -Q weka.classifiers.bayes. net.search.local.K2 -- -P 2 -E weka.classifiers.bayes.net. estimate.SimpleEstimator -- -A 0.5
- Relation: Book1-weka.filters.supervised.attribute.AttributeSelection-Eweka.attributeSelection.CfsSubsetEval-Sweka .attributeSelection.BestFirst -D 1 -N 5
- Instances: 91
- Attributes: 25

 ${\tt PresumptiveDx}$

SxWtloss

SxJaun

SxNau

SxFati

SxPru

SxOT

CxDiab

CTNodeOmit

EUSVascOmit

EUSPortal

EUSNoNode

EUSStagingT

EUSCyto

TxLap

TxRadia

TxChemoTxChemoGemResPxTypeResTransfusionResP0CourseResPathNSurOncNameRadOncNameHistologyTest mode:10-fold cross-validation

=== Classifier model (full training set) ===

Bayes Network Classifier

not using ADTree

#attributes=25 #classindex=24

Network structure (nodes followed by parents)

PresumptiveDx(6): Histology

SxWtloss(2): Histology

SxJaun(2): Histology

SxNau(2): Histology

SxFati(2): Histology

SxPru(2): Histology SxJaun

SxOT(2): Histology

CxDiab(2): Histology SxNau

CTNodeOmit(2): Histology SxPru

EUSVascOmit(2): Histology SxOT

EUSPortal(2): Histology EUSNoNode(2): Histology PresumptiveDx EUSStagingT(5): Histology EUSPortal EUSCyto(7): Histology TxLap(2): Histology SxJaun TxRadia(2): Histology EUSStagingT TxChemo(2): Histology TxRadia TxChemoGem(2): Histology TxChemo ResPxType(7): Histology ResTransfusion(2): Histology CxDiab ResPOCourse(2): Histology CTNodeOmit ResPathN(3): Histology SurOncName(3): Histology EUSCyto RadOncName(6): Histology Histology(14): LogScore Bayes: -1735.470575102397 LogScore BDeu: -232.372454610241 LogScore MDL: -5005.723116036344 LogScore ENTROPY: -2418.745189048893 LogScore AIC: -3565.745189048918

Time taken to build model: 0.08 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	46	50.5495 %
Incorrectly Classified Instances	45	49.4505 %
Kappa statistic	0.3721	
Mean absolute error	0.0767	
Root mean squared error	0.2312	
Relative absolute error	64.7016 %	
Root relative squared error	95.6473 %	
Total Number of Instances	91	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure Class
0.5	0.194	0.48	0.5	0.49
				Adenocarcinoma_of_Pancreas/NOS
0.444	0.037	0.571	0.444	0.5
				Ampullary_Adenocarcinoma
0.667	0.234	0.545	0.667	0.6
				Ductal_Adeno_of_Pancreas
0.4	0.047	0.333	0.4	0.364
				Neuroendocrine_(Islet_Cell)
0	0	0	0	0
				Von_Hippel-Lindau_Syndrome
0	0.011	0	0	0
				Duodenal_Adenocarcinoma
0	0	0	0	0
				Distal_Cholangiocarcinoma
0	0.011	0	0	0

				Renal_Mets
0.25	0.034	0.25	0.25	0.25
				Cystadenoma
0	0	0	0	0
				MEN-I
0	0	0	0	0
				Pseudopapillary_Tumor
0.818	0.063	0.643	0.818	0.72
				IPMN/IPMTBenign_or_CiS
0	0	0	0	0
				Mucinous_Cystic_Neoplasm
0	0	0	0	0
				Benign_Cyst

=== Confusion Matrix ===

a	b	С	d	е	f	g	h	i	j	k	1	m	n	< classified as
12	2	9	1	0	0	0	0	0	0	0	0	0	0	a = Adenocarcinoma_of_Pan
4	4	1	0	0	0	0	0	0	0	0	0	0	0	<pre>b = Ampullary_Adenocarcinoma</pre>
8	0	18	0	0	0	0	0	0	0	0	1	0	0	<pre>c = Ductal_Adeno_of_Pancreas</pre>
1	0	1	2	0	0	0	1	0	0	0	0	0	0	<pre>d = Neuroendocrine_(Islet)</pre>
0	0	0	1	0	0	0	0	0	0	0	0	0	0	e = Von_Hippel-Lindau_Syn
0	1	1	0	0	0	0	0	0	0	0	0	0	0	f = Duodenal_Adenocarcinoma
0	0	0	0	0	1	0	0	0	0	0	0	0	0	g = Distal_Cholangiocarcinoma
0	0	0	1	0	0	0	0	1	0	0	1	0	0	h = Renal_Mets
0	0	1	0	0	0	0	0	1	0	0	2	0	0	i = Cystadenoma
0	0	0	0	0	0	0	0	0	0	0	1	0	0	j = MEN-I

0	0	0	0	0	0	0	0	1	0	0	0	0	0	<pre>k = Pseudopapillary_Tumor</pre>
0	0	1	1	0	0	0	0	0	0	0	9	0	0	<pre>l = IPMN/IPMTBenign_or_CiS</pre>
0	0	1	0	0	0	0	0	0	0	0	0	0	0	<pre>m = Mucinous_Cystic_Neoplasm</pre>
0	0	0	0	0	0	0	0	1	0	0	0	0	0	n = Benign_Cyst

6.2 Classification - Survival - Data Set C - Bayesian Net 2-Parent

Here we have a highly accurate Bayesian Net 2-Parent classifier for survival. This model is taken from the C10 experiments in Section 5.10. Survival prediction is one of the most important topics in oncology research, and is subject of many other research papers (ref. Section 7). As many of these papers use traditional regression methods for survival prediction, it is particularly important here to demonstrate higher performance of novel methods.

A graphical representation of this Bayes Net model is illustrated in Figures 271 and 272. Overall accuracy for this model is rated 60.00%, as compared to average logistic regression performance 42.50% (ref. Figure 228). The accuracy of 60.00% for this single generation of the model exceeds the average iterated performance of the models in C10, which means it outperforms logistic regression via t-testing. There is fairly even coverage across predictions of different survival categories, as shown via the Confusion Matrix in Figure 273.

An interesting feature of this model is the 19 attribute subset chosen via feature-selection. The attributes chosen by feature-selection here contain many elements (diabetes, smoking history, prior chemotherapy treatments, need for palliative measures, etc.) which are known to be highly important in traditional medical assessment of pancreatic cancer survival rates [VD93]. The descriptions of these attribute fields are shown in Table 25. This selection of biologically-correlated attributes makes a strong argument for the medical applicability of this model.

Field	Description
PresumptiveDx	Presumptive Diagnosis
SxSatiety	Presentation - Early Satiety
SxOT	Presentation - Other
CxDiabDiet	Comorbidities - Diabetes Diet Controlled
CxPriorCancerChemo	Comorbidities - Prior Chemo Treatment
SHCigarette	Social History - Cigarettes
PTCDx	PTC Diagnosis
EUSDx	EUS Diagnosis
EUSSMV	EUS - SMV Involvement
EUSNoNode	EUS - No Nodal Involvement
Histology	Histology
PreOutlook	Preliminary Outlook
TxChemoIri	Treatment - Chemotherapy - Irinotecan
TxChemoTax	Treatment - Chemotherapy - Taxol
TxPal	Treatment - Palliation
TxPalStens	Treatment - Palliation - Stenting
ResPOPulmComp	Resection - Postoperative Course - Pulmonary Complications
NoResNoHandle	No Resection - Patient Can't Handle
SurOncName	Surgical Oncologist

Table 25: Survival Feature-Selected Attribute Subset



Figure 271: Classification - Survival - Data Set C - Bayesian Net 2-Parent



Figure 272: Classification - Survival - Data Set C - Bayesian Net 2-Parent (continued)

```
=== Confusion Matrix ===
a b c <-- classified as
14 3 3 | a = '(-inf-5.753425]'
4 11 5 | b = '(5.753425-11.769863]'
4 5 11 | c = '(11.769863-inf)'</pre>
```

Figure 273: Classification - Survival - Data Set C - Confusion Matrix

Weka Output:

=== Run information ===

- Scheme: weka.classifiers.bayes.BayesNet -D -Q weka.classifiers .bayes.net.search.local.K2 -- -P 2 -E weka.classifiers .bayes.net.estimate.SimpleEstimator -- -A 0.5
- Relation: Book1-weka.filters.unsupervised.attribute.Discretize-F-B3-M-1.0-R191-weka.filters.unsupervised.attribute. Remove-R184-weka.filters.supervised.attribute. AttributeSelection-Eweka.attributeSelection .CfsSubsetEval-Sweka.attributeSelection.BestFirst -D 1 -N 5
- Instances: 60
- Attributes: 20

PresumptiveDx

SxSatiety

SxOT

CxDiabDiet

 ${\tt CxPriorCancerChemo}$

SHCigarette

PTCDx

EUSDx

EUSSMV

EUSNoNode

Histology

 ${\tt PreOutlook}$

TxChemoIri

TxChemoTax

TxPal TxPalStens ResPOPulmComp NoResNoHandle SurOncName Longev Test mode: 10-fold cross-validation === Classifier model (full training set) === Bayes Network Classifier not using ADTree #attributes=20 #classindex=19 Network structure (nodes followed by parents) PresumptiveDx(6): Longev

SxOT(2): Longev

CxDiabDiet(2): Longev

CxPriorCancerChemo(2): Longev

SHCigarette(2): Longev SxSatiety

PTCDx(2): Longev

EUSDx(2): Longev

EUSSMV(2): Longev

EUSNoNode(2): Longev PTCDx

Histology(11): Longev PresumptiveDx

PreOutlook(3): Longev EUSSMV

TxChemoIri(2): Longev EUSDx

TxChemoTax(2): Longev
TxPal(2): Longev PreOutlook
TxPalStens(2): Longev TxPal
ResPOPulmComp(2): Longev PTCDx
NoResNoHandle(2): Longev TxPal
SurOncName(3): Longev CxDiabDiet
Longev(3):
LogScore Bayes: -648.3033238760419
LogScore BDeu: -199.86101028520844
LogScore MDL: -1466.702725061748
LogScore ENTROPY: -873.0227635395436
LogScore AIC: -1163.0227635395436

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	36	60	%
Incorrectly Classified Instances	24	40	%
Kappa statistic	0.4		
Mean absolute error	0.3055		
Root mean squared error	0.4237		
Relative absolute error	68.7481 %		
Root relative squared error	89.8773 %		
Total Number of Instances	60		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.7	0.2	0.636	0.7	0.667	'(-inf-5.753425]'
0.55	0.2	0.579	0.55	0.564	'(5.753425-11.769863]'
0.55	0.2	0.579	0.55	0.564	'(11.769863-inf)'

=== Confusion Matrix ===

- a b c <-- classified as
- 14 3 3 | a = '(-inf-5.753425]'
- 4 11 5 | b = '(5.753425-11.769863]'
- 4 5 11 | c = '(11.769863-inf)'

6.3 Regression - ECOG 6-Month - Data Set F - Linear Regression w/ Bagging

Here we have a highly accurate Linear Regression w/ Bagging regressor for 6-Month ECOG scores. This model is taken from the R2 experiments in Section 5.12. This is one of the first experiments where meta-learning affects a statistical improvement to a model. The r-squared value for this model is 0.32, as opposed to 0.26 for standard linear regression, a statistically significant improvement via t-testing. This is also one of the first experiments where machine learning successfully amplifies a traditional predictive regression.

Figure 274 illustrates the Bagging 'committee' which constitutes this model. Each committee member is trained on an N/10 resample of the data set. Training on the resample produces a unique linear regression equation for each member. Each equation uses different coefficients and combinations of attributes from the feature-selected data set. When evaluating a new instance, each member in the committee evaluates and 'votes' on a possible value for 6-Month ECOG. The votes are weighted equally by the model, and an aggregate ECOG prediction is produced. Refer to Section 4.3.5 or [Bre96] for further details on Bagging.

As with most experiments, feature-selected data sets in 6-Month ECOG generally produced more accurate results. Feature-selection generated a 17 attribute subset for data sets F in these experiments. The field names and their explanations are listed in Table 26. Interesting, the majority of these fields involve of chemo regimen and details pertaining to whether a patient underwent resection. These are interesting results, considering that many of these treatment decisions are made directly regarding a patient's potential wellbeing performance.

Field	Description
SxChola	Presentation - Cholangitis
SxBC	Presentation - Biliary Colic
CxDiab	Comorbidities - Diabetes
CxPriorCancerChemo	Comorbidities - Prior Chemo Treatment
EUSDx	EUS Diagnosis
EUSSMVClass	EUS - SMV Involvement Class
EUSCeliacNode	EUS - Celiac Nodal Involvement
ERCPStentType	ERCP Stent Type
TxChemoAva	Treatment - Chemotherapy - Avastin
TxChemoCap	Treatment - Chemotherapy - Capecitabine
TxChemoTax	Treatment - Chemotherapy - Taxol
ResOrgans	Resection - Additional Organs
ResPOAbdominal	Resection - Postoperative Course - Abdominal Collection
ResPOPulmComp	Resection - Postoperative Course - Pulmonary Complications
NoResNoHandle	No Resection - Patient Can't Handle
NoResRefused	No Resection - Patient Refused Treatment
NoResPVInvolve	No Resection - Portal Vein Involvement

Table 26: ECOG 6-Month Feature-Selected Attribute Subset





Weka Output:

=== Run information ===

Scheme:	weka.classifiers.meta.Bagging -P 100 -S 1 -I 10 -W	
	weka.classifiers	
	.functions.LinearRegressionS 0 -R 1.0E-8	
Relation:	Book1-weka.filters.supervised.attribute.AttributeSelection	
	-Eweka.attributeSelection.CfsSubsetEval-Sweka.attribute	
	Selection.BestFirst -D 1 -N 5	
Instances:	72	
Attributes:	18	
	SxChola	
	SxBC	
	CxDiab	
	CxPriorCancerChemo	
	EUSDx	
	EUSSMVClass	
	EUSCeliacNode	
	ERCPStentType	
	TxChemoAVA	
	TxChemoCap	
	TxChemoTax	
	ResOrgans	
	ResPOAbdominal	
	ResPOPulmComp	
	NoResNoHandle	
	NoResRefused	

NoResPVInvolve

ECOG

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

All the base classifiers:

Linear Regression Model

ECOG =

1.0084 * CxDiab=TRUE +
0.6853 * EUSDx=FALSE +
0.9238 * EUSSMVClass=Encased +
1.6886 * EUSCeliacNode=TRUE +
0.5635 * ResPOPulmComp=FALSE +
0.7205 * NoResPVInvolve=TRUE +
-0.4369

Linear Regression Model

ECOG =

0.8263 * CxDiab=TRUE +

0.8083 * EUSDx=FALSE +

1.7504 * EUSCeliacNode=TRUE +

0.8035 * ERCPStentType=Metal +

- 0.6628 * ResPOPulmComp=FALSE +
- 0.6595 * NoResPVInvolve=TRUE +

-0.5322

Linear Regression Model

ECOG =

0.7708 * CxDiab=TRUE + 0.7257 * EUSDx=FALSE + 1.4554 * EUSSMVClass=Encased + 1.0454 * ResPOPulmComp=FALSE + -0.7265 * NoResNoHandle=TRUE + 0.9548 * NoResRefused=TRUE + 1.6817 * NoResPVInvolve=TRUE + -1.4986

Linear Regression Model

ECOG =

0.6748 * CxDiab=TRUE +

```
0.7406 * EUSDx=FALSE +
0.4859 * ResPOPulmComp=FALSE +
0.8273 * NoResPVInvolve=TRUE +
-0.0964
```

Linear Regression Model

ECOG =

1.2688 * SxChola=TRUE +
0.6268 * CxDiab=TRUE +
0.514 * EUSDx=FALSE +
1.3384 * EUSSMVClass=Encased +
1.561 * EUSCeliacNode=TRUE +
0.4906 * NoResRefused=TRUE +
0.1044

Linear Regression Model

ECOG =

0.6175 * CxDiab=TRUE +
0.6021 * EUSDx=FALSE +
1.3825 * EUSSMVClass=Encased +
1.556 * EUSCeliacNode=TRUE +

0.5224 * ResOrgans=spleen +

0.7066 * ResPOPulmComp=FALSE +

1.0307 * NoResRefused=TRUE +

-1.2459

Linear Regression Model

ECOG =

- 0.8249 * SxChola=TRUE +
- 0.7269 * CxDiab=TRUE +
- 0.6689 * EUSDx=FALSE +
- 1.3587 * EUSSMVClass=Encased +
- 1.7183 * EUSCeliacNode=TRUE +
- 0.6451 * ResPOPulmComp=FALSE +
- 0.471 * NoResRefused=TRUE +
- 0.6194 * NoResPVInvolve=TRUE +

-0.8164

Linear Regression Model

ECOG =

0.9979 * SxChola=TRUE + 0.683 * CxDiab=TRUE + 0.7269 * EUSDx=FALSE +

1.4165 * EUSSMVClass=Encased +

0.6263 * NoResRefused=TRUE +

0.5182 * NoResPVInvolve=TRUE +

-0.1991

Linear Regression Model

ECOG =

0.6933 * CxDiab=TRUE +
0.6812 * EUSDx=FALSE +
1.6045 * EUSCeliacNode=TRUE +
0.7885 * NoResPVInvolve=TRUE +
0.3955

Linear Regression Model

ECOG =

0.7232 * CxDiab=TRUE +
0.7302 * EUSDx=FALSE +
1.5916 * EUSCeliacNode=TRUE +
0.8539 * ERCPStentType=Metal +
0.5216 * ResOrgans=spleen +

```
0.6263 * ResPOPulmComp=FALSE +
0.9463 * NoResPVInvolve=TRUE +
-0.8171
```

Time taken to build model: 0.11 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient	0.5706
Mean absolute error	0.4522
Root mean squared error	0.5616
Relative absolute error	72.7155 %
Root relative squared error	81.3339 %
Total Number of Instances	72

6.4 Regression - ECOG 9-Month - Data Set F - Multi-layer Perceptron w/ 2 Hidden Layers

Here we have a highly accurate Multi-layer Perceptron regressor for 9-Month ECOG scores. This model is taken from the R3 experiments in Section 5.13. The r-squared value for this model is 0.16, as opposed to 0.04 for standard linear regression, a statistically significant improvement via t-testing. Multi-layer perceptrons exhibited high r-squared values for many of the regression experiments. They are generally known in medical data mining for high accuracy, but it is difficult to discern from their internal structure how their decisions are produced [KK95].

Figure 275 shows the network layout of this particular regressor. Weights are conditioned via backpropagation on the training sets. Two hidden layers are used, with a learning weight of 0.3 and momentum of 0.2. For new instances, input nodes pass attribute values through the two trained hidden layers, which are aggregated down to produce a ECOG 9-Month prediction. Following the MLP figure is Weka output showing the trained weights on each network connection.

Feature-selected data sets in 9-Month ECOG generally produced more accurate results. Feature-selection generated a 19 attribute subset for data sets F in these experiments. The field names and their explanations are listed in Table 27. As with 6-Month ECOG, the majority of these fields involve chemo regimen and details pertaining the patient's resection. It may be interesting future work to examine whether there is research precedence that these factors significantly affect wellbeing performance.

Field	Description
SxChola	Presentation - Cholangitis
SxBack	Presentation - Back Pain
SxDyspha	Presentation - Dysphasia
CxDiabDiet	Comorbidities - Diabetes Diet Control
CxHyper	Comorbidities - Hypertension
CxPriorCancerChemo	Comorbidities - Prior Chemo Treatment
CXRDx	Chest X-Ray Diagnosis
EUSSMVClass	EUS - SMV Involvement Class
EUSPortal	EUS - Portal Vein Involvement
EUSPortalClass	EUS - Portal Vein Involvement Class
TxChemoAva	Treatment - Chemotherapy - Avastin
TxChemoIri	Treatment - Chemotherapy - Irinotecan
TxChemoLeu	Treatment - Chemotherapy - Leukovorin
TxChemoTax	Treatment - Chemotherapy - Taxol
ResTFFP	Resection - Transfusion - Fresh Frozen Plasma
ResPOLeak	Resection - Postoperative Course - Leak
ResPOAbdominal	Resection - Postoperative Course - Abdominal Collection
ResPOPulmComp	Resection - Postoperative Course - Pulmonary Complications
ResPathR	Resection - Pathology R-Stage

Table 27: ECOG 9-Month Feature-Selected Attribute Subset



Figure 275: Regression - ECOG 9-Month - Data Set F - Multi-layer Perceptron w/2Hidden Layers
Weka Output:

```
=== Run information ===
```

Scheme:	weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2
	-N 500 -V 0 -S 0 -E 20 -H 2
Relation:	Book1-weka.filters.supervised.attribute.AttributeSelection-
	Eweka.attributeSelection.CfsSubsetEval-Sweka.
	attributeSelection.BestFirst -D 1 -N 5
Instances:	72
Attributes:	18
	SxChola
	SxBC
	CxDiab
	CxPriorCancerChemo
	EUSDx
	EUSSMVClass
	EUSCeliacNode
	ERCPStentType
	TxChemoAVA
	TxChemoCap
	TxChemoTax
	ResOrgans
	ResPOAbdominal
	ResPOPulmComp
	NoResNoHandle
	NoResRefused
	NoResPVInvolve

ECOG

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Linear Node 0

Inputs Weights Threshold -0.27907681180500077 Node 1 -0.6689299459785472 Node 2 2.0634974138324895

Sigmoid Node 1

- Inputs Weights
- Threshold 0.2161565560886697
- Attrib SxChola -0.8240512834848186
- Attrib SxBC 0.9719889319602957
- Attrib CxDiab 2.4018587825692728
- Attrib CxPriorCancerChemo 1.5612434307033178
- Attrib EUSDx -1.7402535693664936
- Attrib EUSSMVClass -2.380758092011916
- Attrib EUSCeliacNode -2.2846632584637074
- Attrib ERCPStentType 0.27077130269695104
- Attrib TxChemoAVA 1.6004461214131707
- Attrib TxChemoCap 1.0846704053365328
- Attrib TxChemoTax 1.3085182989378599
- Attrib ResOrgans=spleen -3.775846082844799
- Attrib ResOrgans=duodenum_preserving -2.308761590318322
- Attrib ResOrgans=pylorus-sparing -1.0989356680294642

Attrib	ResPOAbdominal 1.748522222375152		
Attrib	ResPOPulmComp 4.995927948745733		
Attrib	NoResNoHandle -0.8438995490964636		
Attrib	NoResRefused -1.1648150708131346		
Attrib	NoResPVInvolve -1.9970562590263103		
Sigmoid Node 2			
Inputs	Weights		
Thresho	old 0.01931248644125251		
Attrib	SxChola 1.3618203056809814		
Attrib	SxBC -9.480039587078225E-4		
Attrib	CxDiab -2.3505528703455143		
Attrib	CxPriorCancerChemo -0.8645973811683567		
Attrib	EUSDx 2.3745261715454515		
Attrib	EUSSMVClass 3.5097969956607966		
Attrib	EUSCeliacNode 2.617837343077885		
Attrib	ERCPStentType 0.8113622216035797		
Attrib	TxChemoAVA -0.5206862062795482		
Attrib	TxChemoCap 0.12298147304095629		
Attrib	TxChemoTax -0.40371463564664195		
Attrib	ResOrgans=spleen 0.8047396124851516		
Attrib	ResOrgans=duodenum_preserving -0.27889286068420516		
Attrib	ResOrgans=pylorus-sparing -0.20715130545850244		
Attrib	ResPOAbdominal -0.16035164493704002		
Attrib	ResPOPulmComp -2.2285130411821803		
Attrib	NoResNoHandle -0.29341318666906574		
Attrib	NoResRefused 0.5048077573826775		
Attrib	NoResPVInvolve 2.006560027984653		

Class

Input

Node O

Time taken to build model: 0.27 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient	0.4798
Mean absolute error	0.5168
Root mean squared error	0.655
Relative absolute error	83.1096 %
Root relative squared error	94.8561 %
Total Number of Instances	72

7 Related Work

A significant amount of work in medical diagnosis using machine learning has come from the University of Ljubljana, Slovenia, under Prof. Igor Kononenko. [KK95] provides an excellent overview of the medical applicability of machine learning techniques, and presents the advantages and disadvantages of different algorithmic approaches. [Kon93] covers similar ground and presents inductive and Bayesian learning technical for medical analysis in more detail. The techniques discussed in his works have been applied in many medical fields, including pathology, urology, cardiology, and neuropsychology. Work done in [KBK⁺97] applies specifically to oncology, using machine learning to predict the survival time of patients with thyroid carcinoma. The algorithmic focus of this work deals primarily with regression, Assistant decision trees, and Bayesian techniques. We present a broader variety of predictive algorithms in our oncological analysis, and examine different ways to improve algorithmic accuracy, including feature selection and meta-learning.

Machine learning techniques, particularly regression methods, are used commonly in medical literature. [FS03] uses multivariate logistic regression and Cox's proportional hazard model to show that liver metastatis and peritoneal implants are major predictive factors in pancreatic cancer survival. [SR02] contends, using Kaplan-Meier survival analysis, that tumor grading, angioinvasion and perineural invasion are not sufficient pancreatic cancer survival factors. Dr. Murray Brennan makes prolific use of machine learning techniques in his research, and presents in [Bre04] a predictive nomogram for pancreatic cancer survival. Dr. Jennifer Tseng in [Tse04] uses multivariate regression to study survival rates of pancreatic cancer who undergo superior mesenteric or portal vein resections. Our research differs in our broader variety of predictive techniques, and that we look additionally at patient wellbeing and tumor pathology characteristics.

8 Conclusions and Future Work

This thesis set out with two goals-to develop detailed clinical databases of cancer patients, and to conduct machine learning studies on the patient data. With the help of medical professionals at UMass Memorial Hospital, we were able to successfully build clinical databases of seven different cancer forms which can represent the broad narrative of patient treatment. This database was tested by accumulating about a hundred detailed pancreatic cancer patient records. Using this data, we tested a variety of novel machine learning techniques to form predictive models for clinical patient outlook. The accuracy of these novel techniques were statistically tested against linear and logistic regression, the standard medical prediction methods.

We found that most novel machine learning techniques that we tested were able to deliver comparable performance. Both classification and regression algorithms were considered. Generally, Multi-Layer Perceptrons, Bayesian methods, and Locally Weighted Learning with Naive Bayes performed best. In most cases, the novel models performed as well as traditional regression; in some instances they performed even better. Novel regression techniques delivered better performance more frequently than classification techniques. Models based on data sets which used feature selection and supervised discretization generally delivered higher accuracy. In most cases, meta-learning did not improve the accuracy of predictive models. This is a somewhat surprising result, since meta-learning is designed to overcome data mining limitations of smaller data sets.

Future work will expand upon the research basis presented here, and should consider some of the limitations we encountered. First and foremost is attaining a larger patient data set–whether through accumulating additional UMass patients, or expanding the study to include additional institutions or research databases like the HCUP National Inpatient Sample. Continuing to add detail and functionality to the clinical databases will allow for more thorough studies. New knowledge may be gained in testing a populated database module for other gastrointestinal cancers or breast cancer. Studies may be conducted on the individual modules, and clinical performance may even be tested across different disease forms.

There are a broad variety of machine learning predictive algorithms which we did not cover, as well as potential parameter variation for those algorithms we used. There is also algorithm evaluation to consider. In most cases, our novel classifiers had much higher accuracy than logistic regression. However, very often the classifiers performed only as well as a ZeroR guess. The way that single target class values dominate these medical data sets lends itself to predictions for most common class type. This shows simple measurements of accuracy may not always be the best metric of predictive model quality. Other means of evaluation may be necessary and should be explored. Furthermore, the algorithms covered here were based on target class prediction; machine learning to mine association rules and instance clustering has not yet been considered.

The next step in this research should be to continue adding pancreatic patients to the clinical database and generating new predictive models. An informal goal set by Dr. Whalen was to eventually attain classification accuracies of 70% and r-squared values of .50, which makes it clear that more data and further model refinements are still needed. It is important to see whether our experimental results hold up or improve across a broader study population. From the clinical database side, the remaining modules will need further testing and developing. Accumulating clinical data is a critical part of illuminating the design of these modules; much of the functionality of the pancreatic module was decided upon as patient data was being entered and research needs became clearer. Further experiments with neural network based algorithms (MLP, RFB) should be explored in both classification and regression settings, given their initial accuracy and the broad variety of possible algorithm parameters. In experiments where majority classes dominate (t-stage, malignancy, etc.), over-sampling techniques should be explored to emphasize the importance of correctly representing minority classes. Finally, for the more promising predictive models that we've presented here, their performance should be verified against broader pancreatic cancer patient sets, or distinct patient sets from other institutions. This will allow us to conclude the potential of these models for future medical research publication.

9 Acknowledgements

I would first like to thank the Surgical Oncology staff at UMass Memorial Hospital for their time and effort in contributing to this thesis. Dr. Giles Whalen, Mary Sullivan NP, and Dr. Jennifer Tseng made substantial effort in advising development for the gastrointestinal module, as did Dr. Robert Quinlan, Tracy Hall NP, Dr. Rakhshanda Layeeque, Dr. Giulia Cicchetti, and Dr. Ashraf Khan for the breast module. I would like to thank UMass for their generosity in funding this research.

For the computer science portion of this thesis, I would foremost like to thank my advisor Prof. Carolina Ruiz. Her guidance, knowledge, intensity, and patience were essential to completion of this work. Additionally, I would like to thank Prof. Murali Mani for his thesis reading, Prof. George Heineman of WPI for his help on software engineering elements of the project, and Prof. Sergio Alvarez of Boston College for his advising on machine learning experimental design. My appreciation also goes to Provost Carol Simpson and the WPI Graduate School for the 1st-place Science Division award that this work received at the WPI Graduate Research Appreciation Day (GRAD2006) event.

Finally, I would like to thank my friends and family for their encouragement and support in this work, particularly Mom and Dad, without whom none of this would have been possible.

Glossary

Adenocarcinoma: carcinoma which develops within glandular epithelium which typically behaves in a very malignant fashion, 5

Adjuvant: therapy applied post-surgery, 12

- Ampulla of Vater: dilation in the duodenal wall through which the common bile duct and pancreatic duct empty into the small intestine, 7
- Anastomosis: surgically connecting anatomically separate organs to form a continual channel, 3

Benign: cell growth characterized as not spreading to surrounding tissue, 3

- Biopsy: a small sample of tumor tissue taken to evaluate its histologic composition and malignancy, 6
- Cancer: Diseases resulting from uncontrolled cell growth in regions known as neoplasms or tumors, 3
- Carcinogen: Chemical or physical agents which trigger cancer-causing DNA mutations, 3

Carcinoma: cancers arising from epithelial tissue, 3

Celiac axis: artery which originates in the abdominal aorta below the diaphragm, 10

- Chemotherapy: systemic or localized application of antineoplastic drugs to destroy or retard the development of tumor growth, 3
- Computed axial tomography (CT or CAT): a three-dimensional internal view of a patient using a series of sectional x-rays across a common axis, 6
- Cyst: closed cavities of glandular epithelium where retained secretions are accumulated, and may behave in a benign or malignant fashion, 5

- Distal common bile duct: portion of the excretory passage closest to the duodenum which carries bile from the liver, 7
- Duodenum: upper part of the small intestine, which extends from the lower end of the stomach, 7
- ECOG: Eastern Cooperative Oncology Group (ECOG) score for wellbeing, ranges 0-5, consult Table 2, 4
- Endoscopic ultrasound (EUS): ultrasound study generated by a thin, flexible camera passed through the gastrointestinal tract, 6
- Epithelial: related to the epithelium, a membrane of tissue which lines most internal and external surfaces of the body and organs, 3
- Fine needle aspiration (FNA): a biopsy procedure where a sample of cells is obtained applying suction through a fine needle, 10
- G-Stage: refers to grade or differentiation between tumor cells and surrounding normal cells, ranges from 1 to 4, 5
- Gene counseling: series of DNA tests which establish susceptibility of a patient or their family to certain forms of cancer, 3
- Hepatic artery: artery which originates in the celiac artery and supplies the liver with blood, 10
- Histology: the microscopic structure of tumor tissue, 5
- Immunotherapy: for tumors, experimental protocol which uses vaccination to trigger an immune system response which destroys cancerous cells, 3

In situ: tissue growth confined to the site of origin, 5

- Inferior vena cava: vein formed by the union of two iliac veins that transports blood from the lower limbs and pelvic region, 10
- Intraductal papillary mucinous neoplasms: cystic pancreatic tumors which can progress to cancers (called IPMNs or IPMT's), 7

Invasion: malignant cell growth into local tissue, 3

Islet cell tumors: see neuroendocrine tumors, 7

Jejunum: middle part of the small intestine, starts at the end of the duodenum, 7

L-Stage: refers to tumor invasion into lymphatic vessels, 0 if absent and 1 if present, 5

- Lymph Nodes: small bodies along lymphatic vessels which filter bacteria and foreign bodies, presence of tumorous tissue within regional lymph nodes is an important prognostic factor for cancer, 5
- M-Stage: refers to metastatis to distant organs and is denoted 0 if absent and 1 if present, 5
- Magnetic resonance imaging (MRI): use of magnetic resonance of photons to create a highcontrast density image, 6

Malignant: cell growth characterized as spreading to surrounding or distant tissue, 3

Metastasis: malignant cell growth to distant sites in the body, 3

N-Stage: refers to regional lymph node involvement, ranges from 0 to 3, 5

Neoadjuvant: therapy applied pre-surgery, 12

Neoplasm: a distinct mass in a tissue or organ, 3

- Neuroendocrine tumors: tumors which grow in nervous or endocrine tissue and tend to behave in a more indolent fashion than adenocarcinomas, 5
- Oncology: branch of medicine which deals with the diagnosis and treatment of malignant tumors, 3

Palliation: methods intended to relieve cancer symptoms rather than effect a cure, 3

Pancreas: a long gland which sits behind the stomach and secretes digestive juices into the small intestine and bloodstream, 7

Pancreatic cancer: cancer of the pancreas or periampullary region, 7

Pancreaticoduodenectomy: see Whipple procedure, 12

- Periampullary region: area containing the duodenum, distal common bile duct, and ampulla of Vater, 7
- Portal vein: vein that transports blood from the digestive tract, spleen, pancreas, and gallbladder to the liver, 10
- QoL: quality-of-life scores for wellbeing (also known as Karnofsky scores), consult Table 1, 4
- R-Stage: refers to tumor growth on margins of surgically excised tissue: 0 for clean margins, 1 for microscopic tumor growth, and 2 for gross tumor growth, 5

Radiotherapy: treatments which use irradiation to destroy cancerous cells, 3

Resection: surgical excision of tumor growth from bodily tissue, 3

Serum study: a blood test, which may include nutritional levels, liver functions, and molecular tumor markers, 6 Splenic vein: vein generated from several smaller veins which meet at the front surface of the spleen, 10

Stenting: propping open an anatomical vessel with a metal or plastic stent, 3

- Superior mesenteric artery: artery which originates from the upper aorta which supplies the small intestines and colon, 10
- Superior mesenteric vein: vein which begins at the ileum and joins behind the pancreas with the splenic vein, 10

T-Stage: refers to primary tumor size, ranges from 0 to 4 or 'is' for in situ growth, 5

Tumor markers: molecular systemic indicators of certain cancer forms, 6

Tumor: a distinct mass in a tissue or organ, 3

Ultrasound: use of ultrasonic waves to create a sonographic visualization a body's internal structure, 6

V-Stage: refers to tumor invasion into veins, 0 if absent and 1 if present, 5

Vasculature: blood vessels; penetration of tumors into vasculature can be an important factor in determining the spread and resectability of the disease, 5

Whipple procedure: most common surgical procedure to treat pancreatic cancer, 12

X-ray: the process of visualizing an internal body image by catching high-energy photons on photographic film, 6

References

- [Aka74] Hirotugu Akaike. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19(6), December 1974.
- [AKA91] David W. Aha, Dennis Kibler, and Marc K. Albert. Instance-Based Learning Algorithms. Machine Learning, 6(1):37–66, 1991.
- [AMS97] Chris Atkeson, Andrew Moore, and Stefan Schaal. Locally Weighted Learning. AI Review, 11:11–73, April 1997.
- [Bre96] Leo Breiman. Bagging Predictors. *Machine Learning*, 24(2):123–140, 1996.
- [Bre04] MF Brennan. Prognostic Nomogram for Patients Undergoing Resection for Adenocarcinoma of the Pancreas. Annals of Surgery, 240:1–6, 2004.
- [Cen] Federal Citizen Information Center. Consumer Information Center: Gallstones. http://www.pueblo.gsa.gov/cic_text/health/gallstones/gallstns.htm.
- [Cli] Mayo Clinic. Whipple Procedure. http://www.mayoclinic.org/pancreaticcancer/whippleprocedure.html.
- [DD04] J. DeWitt and B. Devereaux. Comparison of Endoscopic Ultrasonography and Multidetector Computed Tomography for Detecting and Staging Pancreatic Cancer. Annals of Internal Medicine, 141:753–763, 2004.
- [Dev95] Jay L. Devore. Probability and Statistics for Engineering and the Sciences. Brooks/Cole Publishing Company, 4th edition, 1995.
- [FHP03] Eibe Frank, Mark Hall, and Bernhard Pfahringer. Locally Weighted Naive Bayes. In Proceedings of the 19th Annual Conference on Uncertainty in Artificial Intelligence (UAI-03), pages 249–25, San Francisco, CA, 2003. Morgan Kaufmann.

- [FI93] Usama M. Fayyad and Keki B. Irani. Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning. In *IJCAI*, pages 1022–1029, 1993.
- [FS96] Yoav Freund and Robert E. Schapire. Experiments with a New Boosting Algorithm. In *International Conference on Machine Learning*, pages 148–156, 1996.
- [FS03] Y Fujino and Y Suzuki. Predicting Factors for Survival of Patients with Unresectable Pancreatic Cancer: A Management Guideline. *Hepatogastroenterology*, 49:250–253, 2003.
- [Gra95] H. Gray. *Gray's Anatomy*. Barnes and Noble Books, 15th edition, 1995.
- [Hal98] M. Hall. Correlation-Based Feature Selection for Machine Learning. PhD thesis, Waikato University, Department of Computer Science, 1998.
- [Ins] Sanger Institute. Cancer Genome Project. http://www.sanger.ac.uk/genetics/CGP/.
- [IW05] E. Frank I. Witten. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 2nd edition, 2005.
- [KB49] DA Karnofsky and JH Burchenal. The Clinical Evaluation of Chemotherapeutic Agents in Cancer. In Evaluation of Chemotherapeutic Agents, page 196. Columbia Univ Press, 1949.
- [KBK⁺97] M. Kukar, N. Besic, I. Kononenko, M. Auersperg, and M. Robnik-Sikonia. Prognosing the Survival Time of the Patients with the Anaplastic Thyroid Carcinoma with Machine Learning. In *Intelligent Data Analysis in Medicine and Pharma*cology, pages 116–129, 1997.
- [KK95] I. Kononenko and M. Kukar. Machine Learning for Medical Diagnosis: History, State of the Art, and Perspective. In Proc. Workshop on Computer Aided Data Analysis in Medicine and Pharmacology, pages 9–31, Bled, Slovenia, 1995.

- [Kon93] I. Kononenko. Inductive and Bayesian Learning in Medical Diagnosis. Applied Intelligence, 7:317–337, 1993.
- [lCvH92] S. le Cessie and J.C. van Houwelingen. Ridge Estimators in Logistic Regression. Applied Statistics, 41(1):191–201, 1992.
- [MD89] J. Moody and C. Darken. Fast Learning in Networks of Locally-Tuned Processing Units. Neural Computation, 1(2):281–294, 1989.
- [Mit97] T. Mitchell. Machine Learning. McGraw-Hill Science/Engineering/Math, 1st edition, 1997.
- [OC82] MM Oken and RH Creech. Toxicity and Response Criteria of the Eastern Cooperative Oncology Group. American Journal of Clinical Oncology, 5:649–655, 1982.
- [oC04] American Joint Committee on Cancer. Collaborative Staging Manual and Coding Instructions, version 1.0. Jointly published by American Joint Committee on Cancer (Chicago, IL) and U.S. Department of Health and Human Services (Bethesda, MD), 1st edition, 2004.
- [Qui92] J. R. Quinlan. Learning with Continuous Classes. In 5th Australian Joint Conference on Artificial Intelligence, pages 343–348, 1992.
- [Qui93] J. Ross Quinlan. C4.5: Programs for Machine Learning. Morgan Kaufmann, Los Altos, California, 1993.
- [Soc] American Cancer Society. Cancer Facts and Figures 2005. http://www.cancer.org/downloads/STT/CAFF2005f4PWSecured.pdf.
- [SR02] R. Strnad and M. Ryska. Are We Able to Predict Survival Rate after the Radical Resection of the Pancreas for the Pancreatic Ductal Adenocarcinoma? In *Joint*

Meeting of the European Pancreatic Club (EPC) and the International Association of Pancreatology (IAP), Heidelberg, Germany, June 2002.

- [Szo82] P. Szolovits. Artificial Intelligence and Medicine. Westview Press, Boulder, Colorado, 1982.
- [Tse04] Jennifer F. Tseng. Resection Of The Superior Mesenteric-Portal Vein For Pancreatic Adenocarcinoma: Margin Status And Survival Duration. In Proceedings of the 45th Annual Meeting of the Society for Surgery of the Alimentary Tract, New Orleans, LA, May 2004.
- [VD93] S. Rosenberg V. DeVita, S. Hellman. Cancer: Principles and Practices of Oncology. J. B. Lippincott, 4th edition, 1993.
- [Wik] Wikipedia: The Free Encyclopedia. http://en.wikipedia.org/wiki/Main_Page.
- [Wol90] D. H. Wolpert. Stacked Generalization. Technical Report LA-UR-90-3460, Los Alamos National Laboratory, Los Alamos, NM, 1990.