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	ut 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	11
4	Pancreatic Cancer N-Staging	11
5	Pancreatic Cancer M-Staging	11
6	Tumor Size Distribution	115
7	T-Stage Distribution	115
8	N-Stage Distribution	115
9	Vasculature Involvement Distribution	115
10	Histology Distribution	116
11	Malignancy Distribution	116
12	ECOG 6-Month Distribution	116
13	ECOG 9-Month Distribution	116

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 *neoplasms* 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		
ΤХ	Primary tumor cannot be assessed		
Τ0	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
MO	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

Eile Edit Yiew Insert	Format <u>R</u> ecords <u>T</u> ools	Window Help Adobe	PDF	_ 8
4-8860.*	8 B B 0 8 2	31 V Ta V A	• • • • 🖻 🗗 1	· 2.
MS Sans Seri	f 🔸 8 🔸 🖪	<i>I</i> <u>U</u> ≡ ≡ ≡	<u>⊘</u> • <u>∧</u> • <u>⊿</u> •	
resumptive Diagr	osis at Onset of	Care		
Sec. A				
resentation				
resentation Date of Evaluation ECOD	Performance Status	Height (in.) Weig	<u>ht (lbs.)</u>	1
Image: mail of the second and the	<u>i Performance Status</u> 何1 何2 何3 何4	Height (in.) Weig	<u>ht (lbs.)</u>]
Image: matrix of Evaluation ECOC Date of Evaluation © 0 Symptoms	i <u>Performance Status</u> ぼ1 ぼ2 ぼ3 ぼ4	Height (in.) Weig	<u>ht (lbs.)</u>]
Image: matrix of the second and th	<u>i Performance Status</u> ☞ 1 ☞ 2 ☞ 3 ☞ 4 庫 Biliary Colic	Height (in.) Weig	ht (Ibs.)	
Image: constraint of the second and the second an	i Performance Status @ 1 @ 2 @ 3 @ 4 @ Biliary Colic @ Nausea	Height (in.) Weig	ht (lbs.) Indigestion Dysphagia	
Tesentation Date of Evaluation ECOU Image: Symptoms Image: Symptoms Image: Weight Loss How Much (pounds): Image: How Much (pounds): Image: Symptoms	i <u>Performance Status</u> 1 1 2 1 3 4 II Biliary Colic II Nausea IV Vomiting	Height (in.) Weig	ht (lbs.) Indigestion Dysphagia Early Satiety	
Image: constraint of the second and the second an	i Performance Status 1 2 3 4 Biliary Colic Nausea Vomiting Clay Colored Stool	Height (in.) Weig	ht (lbs.) Indigestion Dysphagia Early Satiety	

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
5xChola	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
5xOT	Yes/No	Initial Symptoms - Other
SxOTSpe	Text	Initial Symptoms - Other - Specify
	11086941	

Figure 6: Pancreatic Cancer Presentation Schema

Microsoft Access - [Pan_2	_History : Form]	A DESCRIPTION OF THE OWNER OF THE OWNER OF THE	-OX
🗄 File Edit View Insert	: Format <u>R</u> ecords <u>T</u> ools !	Window Help Adobe PDF	_ 8 ×
⊻ • 8 8 8 8 *			2.
Medical History			-
Comorbidities Heart Failure Heart Failure Respiratory Renal Failure Hypertension Bleeding Disorder Social History Cigarette Use Respiratory Respiratory Respiratory Cigarette Use Respiratory Respirat	Malnutrition Liver Failure/Cirrhosis Diabetes Circles than Six Months Gireater than Six Months	Cancer History Patient Prior Dx: Im Chemo Im Chemo <th></th>	
Form View			

Figure 7: Pancreatic Cancer Medical History Form

ġ,	Field Name	Data Type	Description
8)D	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
Ĩ.	CxHF	Yes/No	Comorbidities - Heart Failure
8	CXIHD	Yes/No	Comorbidities - Ischemic Heart Disease
	CxResp	Yes/No	Comorbidities - Respiratory
	CxDiab	Yes/No	Comorbidities - Diabetes
Ũ.	CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral
ŝ	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
1	CxHyper	Yes/No	Comorbidities - Hypertension
8	CxBleed	Yes/No	Comorbidities - Bleeding Disorder
	CxLiver	Yes/No	Comorbidities - Liver Failure
	CxMal	Yes/No	Comorbidities - Malnutrition
1	CxPriorCancer	Number	Comorbidities - Prior Cancer Dx
8	CxPriorCancerChemo	Yes/No	Comorbidities - Prior Cancer Dx - Chemo
	CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation
0	CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery
1	SHCigarette	Yes/No	Social History - Cigarettes (significant use)
8	SHAlcohol	Yes/No	Social History - Alcohol (significant use)
	SHDrugUse	Yes/No	Social History - Drug Use
0	SHExposure	Yes/No	Social History - Environmental Exposure
1	SHOther	Yes/No	Social History - Other
8	SHOtherS	Text	Social History - Other - Specify
	FamilyFatherDx	Number	Family History - Father Dx
	FamilyMotherDx	Number	Family History - Mother Dx
1	FamilyOther1	Text	Family History - Other1
ŝ	FamilyOther1Dx	Number	Family History - Other1 Dx
	FamilyOther2	Text	Family History - Other2
U	FamilyOther2Dx	Number	Family History - Other2 Dx

Figure 8: Pancreatic Cancer Medical History Table Schema
E File	Edit	⊻iew	Insert	Format	Re	cords	Too	ls !	<u>M</u> indov	N	Help	Ad	lo <u>b</u> e P	DF				- 8	×
X •	🖬 🔁	6	<u>à</u> 🚏	X 🗈	R	ю	6	₹↓	Z.	V/	酒	∇	44	>*	NN I	Ð	•	2.	
		MS	Sans Seri	f	×	8	÷	в	I	U	F	畫	-	۵.	A	4 -		•	×
Serum	Stud	ies															-		
CEA:			Albumir	າ]			AL	.К				A	λLT	Γ		1			
CA19-9:		_	Total B	ilirubin			AS	ST:				A	mylas	:e: [

Figure 9: Pancreatic Cancer Serums Studies Form

335	Field Name	Data Type	Description	
P	ID	AutoNumber	ID	
11°	MR	Text	Meditech Medical Record Number for Patient	
ŝ	LabCEA	Number	Laboratory - CEA	
	LabCA19-9	Number	Laboratory - CA19-9	
	LabAlb	Number	Laboratory - Albumin	
1	LabBili	Number	Laboratory - Bilirubin	
ŝ	LabAlka	Number	Laboratory - Alkaline phosphotase	
	LabALT	Number	Laboratory - ALT	
	LabAST	Number	Laboratory - AST	
1	LabAmylase	Number	Laboratory - Amylase	
8				

Figure 10: Pancreatic Cancer Serums Studies Table Schema

Microsoft Access - [Pan_3b_DiagImg :	Form]	
🗐 File Edit View Insert Format I	tecords Tools Window Help Adobe PDF Type a question for help 👻 .	. 8 ×
₩ • ■ 1 ● 0 ♥ % 1 6	s ∽ < < ; ; ÿ ъ ⊽ M ↦ K 🖻 🗗 + Q .	
	· B Z U ≡ ≡ ■ <u>></u> · <u>A</u> · <u>⊿</u> · □ ·	Į.
Diagnostic Imaging Procedur	es	
CT with Panreatic Protocol/CT/	s	
Date of Procedure:		
🕅 Celiac Artery Involvement	C Open C Abutted C Encased C Occluded C Unknown	
🕅 Superior Mesenteric Artery Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
🔚 Hepatic Artery Involvement	Copen C Abutted C Encased C Occluded C Unknown	
🕅 Inferior Vena Cava Involvement	Copen C Abutted C Encased C Occluded C Unknown	
🕅 Superior Mesenteric Vein Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown	
🕅 Portal Vein Involvement	Ø Open Ø Abutted Ø Encased Ø Occluded Ø Unknown	
Nodes: 🛛 🖾 Celiac Nodal Disease	🕅 Other Nodal Disease 🛛 🕅 No Nodal Assessment or Mention	
Tumor Size (cm): by		
Chest X-Ray (CXR)		
E Parcutanopus Transhonatis Ch	olengiography (PTC)	
Date of Procedure		
Stenting Type Conternal	@ External	
Type Interna		
Form View		1

Figure 11: Pancreatic 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
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 12: Pancreatic Cancer Diagnostic Imaging Table Schema

Microsoft Access - [Pan_3c_Endoscopy	y : Form]	
🗐 File Edit View Insert Format F	Records Tools Window Help Adobe PDF Type a question for help 👻 🔔	8 ×
	• BIUEEEA.A.L.	
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	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
🔟 Inferior Vena Cava Involvement	© Open @ Abutted @ Encased @ Occluded @ Unknown	
🕅 Superior Mesenteric Vein Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknown	
🕅 Portal Vein Involvement	Ø Open Ø Abutted Ø Encased Ø Occluded Ø Unknown	
Nodes: Celiac Nodal Disease Tumor Size (cm): by	Dther Nodal Disease 🖉 No Nodal Assessment or Mention	
EUS Staging: T 💽 N	FNA Cytology	
Endoscopic Retrograde Chola	ingiopancreatogram (ERCP)	
Date of Procedure:		
IIII <u>Stenting</u> Type: [©] Plastic	Metal	
Form View		1

Figure 13: Pancreatic Cancer Endoscopy Studies Form

Field Name	Data Type	Description
80 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

Microsoft Access - [Pan_4_Prelim : Form]	
🗐 File Edit View Insert Format Records	Tools Window Help Adobe PDF Type a question for help 🗸 🕳 🛪
🔟 - 🖬 🕲 🎒 🖪 🖤 👗 🖻 🛍 🗠	A A Y B ∇ A >> × × B □ □ · · Q.
*	· BZU≣≣≣⊉·A·⊿·□·.
Pre-Surgical Outlook	
Potentially Resectable	1
Locally Advanced/Unresectable	
Metastatic or Equivocal Findings	
Form View	

Figure 15: Pancreatic Cancer Preliminary Outlook Form

6	Field Mome	Data Type	Description
8 ID	Search	AutoNumber	ID and an an an an and an an and
MR		Text	Meditech Medical Record Number for Patient
PreC	outlook	Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic or Equivocal Findings)
255			

Figure 16: Pancreatic Cancer Preliminary Outlook Table Schema

EB File Edit View	Insert Format Records	<u>Tools Window H</u> elp Ad	obe PDF Type a question I	iorhelp 🔹 🗕 🗗 >
🔟 • 日 🔁 🖨 [ð 💱 🕺 🖻 💼 🗠		🗚 >* 🕅 🗗 🖆	1 · 🛛 .
	*	• B I U = =	≡ <u>⊅</u> + <u>∧</u> + <u>√</u> + [· · · ·
Freatment Co	<u>urse</u>			
Resection		Staging Laparos	copy/Laparotomy	
Radiation		🗏 Palliative Measu	res	
Adjuvant Neoa	djuvant 🍘 Both	Bypass	T HAL	
Chemotherapy		🕅 Gastrostomy Tube	👿 PV Shunts	
l 🖉 Adjuvant 🖗 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:	Experimental Pro	otocol (vaccine, etc.)	
🕅 Irinotecan		Genetic Counsel	ing	

Figure 17: Pancreatic Cancer Treatment Form

Field Name	Data Type	Description
₿▶ ID	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 - Fluorouracii (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 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 + ■ • ● ● ● * * ■ ■ + • • ● ● # # ♥ • ▼ ● + * * ● ● ● + ₽ .
- MS Sans Serif • 8 • B I U = = = 🖄 • 🗛 • 🖉 • 🗁 •
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:
Eorm 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

Microsoft Access - [Pan_6b_NoRes : Form]	<u>_0×</u>
🗐 File Edit View Insert Format Records	Iools Window Help Adobe PDF Type a question for help 🔹 🕳 🛪
₩ • ■840° × 88 ∞	
	· B I U ≡ ≡ ≡ <u>></u> · <u>A</u> · <u>I</u> · □ · .
If Resection is Not Performed	
Date of Decision:	
Reasons (select all that apply):	
<u>Clinical Decision</u>	
🕅 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	
Form View	

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 Access - [Pan_7_Path : For	rm]		
E File Edit View Insert Format	<u>R</u> ecords <u>T</u> ools <u>W</u> indow <u>H</u> el	Ip Adobe PDF Type a question for help 👻 .	- 8 ×
🔟 • 🖬 🗞 🎒 🖧 🖤 🔏 🖿	n 8 2 X Y Y	ā⊽ 🗚 ↦ ₩ 🖻 🗗 🖬 + 🛛 .	
MS Sans Serif	• 10 • B I <u>U</u> ≣	E ≣ ≣ 🏖 • 🔺 · 🖉 • 🔲 •	
Final Tumor Histology			
[] (from best of imaging, FNA, pathology, etc)			
Pathology (if available)			
Tumor Size (cm): by			
TNM Staging: T	✓ M ✓ R:		
Form View			11.

Figure 23: Pancreatic Cancer Pathology Form

Data Type	Description
AutoNumber	ID
Text	Meditech Medical Record Number for Patient
Number	Histology
Number	Resection - Pathology Staging - T
Number	Resection - Pathology Staging - N
Number	Resection - Pathology Staging - M
Number	Resection - Pathology Staging - R
Number	Resection - Pathology Staging - R
Number	Resection - Pathology Tumor Size (cm) - Width
Number	Resection - Pathology Tumor Size (cm) - Height
	AutoNumber Text Number Number Number Number Number Number Number Number Number

Figure 24: Pancreatic Cancer Pathology Table Schema

Microsoft Access - [Follow-	up Information]	- 🗆 🗵
🗐 File Edit View Insert	Format Records Tools Window Help Adobe PDF Type a question for help:	×
🔟 • 日 🗞 🥔 🗛 🖤	¾ ๒ ඬ ∽ ⊛ ⋬ ᠯ ♥ Ѣ ▽ ₩ ₩ 匝 亩 • ♡.	
Pancreatic Tumor		
Follow-up Informa	tion Add Record Find Record Delete Record	
MR 000726070	Follow-up Window: 9M -	
Date of Visit: 05/11/2	003 Weight (pounds): QOL score:	
ECOG performance status:	0 @ 1 @ 2 @ 3 @ 4	
Lab Value:		
CEA: Alt	umin Alkaline Phosphotase	
CA19-9: To	tal 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	T Fatigue	
Status:		
Died Death Date:	5711/2003	
O NED		
C AWD Method of Detection	v 🗖 Lab 🔲 Badiologic Evidence 🗖 Clinical Evidence	
Form View		1

Figure 25: Pancreatic Cancer Follow-Up Form

Field Name	Data Type	Description
8 ID	AutoNumber	
MR	Text	Meditech Medical Record Number for Patient
FUWin	Number	Follow-Up Windows
VisitDate	Date/Time	Visit Date
Weight	Number	Weight (lbs.)
QOLscore	Number	QoL Score (0-100)
ECOG	Number	ECOG Score (0-4)
LabCEA	Number	Laboratory - CEA
LabCA19-9	Number	Laboratory - CA19-9
LabAlb	Number	Laboratory - Albumin
LabBili	Number	Laboratory - Bilirubin
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
SxChola	Yes/No	Symptoms - Cholangitis
SxBC	Yes/No	Symptoms - Biliary Colic
SxNau	Yes/No	Symptoms - Nausea
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
SxAbd	Yes/No	Symptoms - Abdominal Pain
SxBack	Yes/No	Symptoms - Back Pain
SxOT	Yes/No	Symptoms - Other
SxOTSpe	Text	Symptoms - Other - Specify
Status	Number	Status (NED, AWD, Died)
DeathDate	Date/Time	Death Date
StatusAWDLab	Yes/No	AWD - Lab Evidence
StatusAWDRad	Yes/No	AWD - Radiology Evidence
StatusAWDCli	Yes/No	AWD - Clinical Evidence
	1000000	

Figure 26: Pancreatic Cancer Follow-Up Table Schema

3.1.2 Hepatocellular Cancer

The Fac Ten Direct 1	onnac Records Tools	window Telb Adob		
- 🖬 🛍 🖨 🖾 🖤 🛛	6 00 00 10 18 24	¥† 12 🖉 🛆 🕴	Ma 🕨 📈 😭 🖆	Ì /॑॑ - └ू) -
MS Sans Serif	• 8 • B	<i>I</i> <u>U</u> ≣ ≡ ≡	I 🕸 • 📥 • 🚄	•
esumptive Diagno	sis at Onset of	Care		
ee amp are bragine				
	<u> </u>			
coontation				
esemanon				
ate of Evaluation ECOG P	erformance Status	<u>Height (in.)</u> Weig	ght (lbs.)	
» 0 »	1 @ 2 @ 3 @ 4			
umptome				
<u>Ymptoms</u>				
Weight Loss	🕅 Biliary Colic	🕅 Pruritis	🕅 Indigestion	
Weight Loss How Much (pounds):	🕅 Biliary Colic 🕅 Nausea	🕅 Pruritis 🕅 Abdominal Pain	🕅 Indigestion 🕅 Dysphagia	
Weight Loss How Much (pounds):	I Biliary Colic I II Nausea III Vomiting	IIII Pruritis IIII Abdominal Pain IIIII Back Pain	Indigestion III Dysphagia IIII Early Satiety	
Weight Loss How Much (pounds): Journalice Cholecystitis	I Biliary Colic II Nausea III Vomiting IIII Clay Colored Stool	 Pruritis Abdominal Pain Back Pain Other Specify: 	III Indigestion IIII Dysphagia IIII Early Satiety	
Wight Loss How Much (pounds): Jaundice Cholecystitis Cholangitis	 Im Biliary Colic Im Nausea Im Vomiting Im Clay Colored Stool Im Fatigue 	Prunitis Abdominal Pain Back Pain Other Specify:	Indigestion III Dysphagia III Early Satiety	

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
5xOT	Yes/No	Initial Symptoms - Other
SxOTSpe	Text	Initial Symptoms - Other - Specify
	1008694C	
2		

Figure 28: Hepatocellular Cancer Presentation Table Schema

Microsoft Access - [HCC_2	_History : Form]			
🗐 Eile Edit View Insert	: Format <u>R</u> ecords <u>T</u> ools <u>y</u>	<u>∦indow H</u> elp Ado <u>b</u> e PDF	Type a question for h	ielp 👻 🗕 🗗 🗙
🔟 • 🖬 🖶 🖨 🖪 🖤	X 🖻 🖻 🗠 🍓 🛃	XI V TO V M .	* 🕷 🗗 🗗 🕯 (2).
*	* * B	1 U = = = 2	<u>- A - 2</u> -	□••
Medical History				
<u>Comorbidities</u>	n	Cancer History		
IIII Heart Failure		Patient Prior Dx:	•	
IIII Ischemic Heart Disease	III Liver Failure/Cirrhosis	🕅 Chemo 🕅 Radiatio	on 🖾 Surgery	
III Respiratory	I™ Diabetes	Father Dy:		
🕅 Renal Failure	Less than Six Months Greater than Six Months			
With Hypertension	Cral Agente	Mother Dx:	<u> </u>	
Bleeding Disorder	Time Diet Control	Other Relation:		
Social History		Related Dv:		
🕅 Cigarette Use 🛛 🕅 Irregu	lar Drug Use	Other	<u> </u>	
🕅 Alcohol Use 🛛 🕅 Envir	ormental Exposure	Relation:		
🕅 Other - Specify:		Related Dx:	•	
		<u>1</u>		
Form View				1

Figure 29: Hepatocellular Cancer Medical History Form

Field Name	Data Type	Description
80 D	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
CxPriorCancerRadiat	ion Yes/No	Comorbidities - Prior Cancer Dx - Radiation
CxPriorCancerSurger	y 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

🔚 Eile Edit	<u>V</u> iew Insert	: F <u>o</u> rmat	Rec	ords	Tool	s ⊻	/indov	v H	elp	Ado <u>b</u> e I	PDF		Туре	a que:	stion fo	or help		_ 8 ×
🔟 - 🔛 🖲	a 🛛 🖉	" 洗 暄	8	ю		₹↓	Z	V.	6	7 44	•	120	P		•	2	-	
	- MS Sans Se	erif	F	8	÷	в	I	U	F 1	F 31	2.	A	÷	1 -		• =	· • .	
Serum S	tudies									-			1					
CEA:	Albun	nin [AL	ĸ	Г			ALT	Γ			5	14			
C410.0	Total	Bilirubin [_	AS	T:		_	-	Amyla	ise:	_	_					

Figure 31: Hepatocellular Cancer Serum Studies Form

Field Name	Data Type	Description	
80 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	

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 🚽 🗗 🗙
<u>▶</u> • ■ • ● ● ● ● ● ● ● ● ● ●
MS Sans Serif • 8 • B I U ≣ ≣ ≣ ⊉ • ▲ • ⊿ • □ • •
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
Form View

Figure 33: Hepatocellular Cancer Diagnostic Imaging Form

Field Name	Data Type	Description			
80 (D	AutoNumber	ID			
MR	Text	Meditech Medical Record Number for Patient			
ChildsC	Number	Childs Class			
TStage	Number	Imaging T-Stage			
NStage	Number	aging 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			
PChemo_AVA	Yes/No	Prior Chemo - Avastin			
PChemo_Cap	Yes/No	Prior Chemo - Capecitabine			
PChemo_Erb	Yes/No	Prior Chemo - Erbitux			
PChemo_Flu	Yes/No	Prior Chemo - Fluorouracil (5-FU)			
PChemo_FUDR	Yes/No	Prior Chemo - FUDR			
PChemo_Gem	Yes/No	Prior Chemo - Gemcitabine			
PChemo_Iri	Yes/No	Prior Chemo - Irinotecan			
PChemo_Leu	Yes/No	Prior Chemo - Leukovorin			
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]
🐵 Eile Edit Viev	v Insert Format Records Iools Window Help Adobe PDF 🛛 Type a question for help 🔽 🗗 🗙
	• Q. ♥ % ■ C ↔ @ 21 X ▼ T ▼ T ♥ A > ♥ Ø D = • Q. → B / U = = = 2 • A - 2 • [• □•.
Disease Stag	<u>1e</u>
B or C cirrhosis B or C cirrhosis Non-cirrhotic or Chil	Correction Constraint Candidate Constraint Candidate
Pre-Surgical	Outlook
Potentially Rese	ectable
Locally Advance	ed/Unresectable
G Metastatic or E	quivocal Findings
Pursue Rese	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
SuroTreat	Number	Initial Surgical Decision (Resect, Ablate, Resect/Ablate)

Figure 36: Hepatocellular Cancer Preliminary Outlook Table Schema

Picrosoft Access - [H	ICC_5_Treatment : Form]			-0×
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
¶\$▶1D	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 38: Hepatocellular Cancer Treatment Table Schema

Elle Edit View Insert For ₩ • ₩ ₩ ⊕ ₽ ₩ ₩ ×	mat Records Iools Window Help Adobe PDF Type a question for help • ■ Particle PDF Type a question for help • ■	8×
Ablation		
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	
Form View		

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]
🗉 Eile Edit 💟 Insert Format Records Iools Window Help Adobe PDF 🛛 Type a question for help 🔽 🕳
¥· ₩ 8 8 0 % % 8 8 0 % \$ \$ \$ \$ 7 8 7 8 1× × 8 6 1 4 .
M5 Sans Serif • 8 • B I U ≣ ≣ ≣ 2 • ▲ • 2 • - •
If 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
I Wound Infection I Congruent with Post Ωρ Care Path I NG/Gastrostomy Drainage > 7 Days
Eleak 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:
Metastatic gastric ademo with resectable primary tumor:
🖉 Primary Tumor Grossly Symptomatic Volume of Metastatic Disease: 💽
Form View

Figure 41: Hepatocellular Cancer Resection Form

Jus E	Field Name	Data Type	Description				
8 ID		AutoNumber	ID				
MR		Text	Meditech Medical Record Number for Patient				
DAdn	n	Date/Time	ce of Admission				
DSun	g	Date/Time	Date of Surgery				
ORTI	me	Number	OR Time				
PxTy	pe	Number	Procedure Type				
Orga	ins	Text	Other Organs Resected				
Blood	dLoss	Number	Blood Loss (cc)				
CVP		Yes/No	CVP cm/H20 During Hepatic Transection/Puncture				
CVPc	m	Number	CVP cm/H20				
Tidal	cc	Number	Tidal Volume During Hepatic Transection/Puncture (cc)				
Tran	sfusion	Yes/No	Transfusion - Needed?				
T Ur	nits	Number	Transfusion - Units				
T FF	P	Yes/No	Transfusion - Fresh Frozen Plasma				
T Ce		Yes/No	Transfusion - Cell Saver				
T Tis	sue	Yes/No	Transfusion - Tissue Link				
T CV	/SA	Yes/No	Transfusion - CV5A				
T Fir	naer	Yes/No	Transfusion - Finger Fracture				
T Ck	amp	Yes/No	Transfusion - Clamp/Crush				
T Ar	aon	Yes/No	Transfusion - Ardon Beam				
T Pa	rv	Yes/No	Transfusion - Staplers for Parvnchyma				
T Str	ruct	Yes/No	Transfusion - Staplers for Structures				
T Gh	Je	Yes/No	Transfusion - Fibrin Glue				
T RF	A	Yes/No	Transfusion - RFA				
TCr	v	Yes/No	Transfusion - Cryoablation				
ICUd	lavs	Number	Days in ICU				
Infec	tion	Yes/No	PO - Wound Infection				
Leak		Yes/No	PO - Leak				
Conc	ruent	Yes/No	PO - Congruent Post-Op Path				
Diver	raent	Yes/No	PO - Divergent Post-Op Path				
NG	124	Yes/No	PO - NG/Gastrostomy Drainage > 7 Days				
Abde	ominal	Yes/No	PO - Abdominal Collection				
Pulm/	Cx	Yes/No	PO - Pulmonary Complications				
Cath	eter	Yes/No	PO - Catheter Infections				
Drain	IS	Yes/No	PO - Drains				
Rena	alInsuf	Yes/No	PO - Renal Insufficiency				
Liver	Insuf	Yes/No	PO - Liver Insufficiency				
LI TF	80.00	Number	PO - Liver Insufficiency - Total Bilirubin				
DDisc	charge	Date/Time	Date of Discharge				
Disch	Status	Number	Discharge Status				
▶ DDez	ath	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

PMicrosoft Acc	ess - [HCC_8_Path : Fo	irm]					-O×
E Eile Edit	View Insert Format	Records To	ols <u>W</u> indow	Help Adobe	PDF	Type a question for help	8 ×
🔛 • 🔛 📆	a 🔉 💱 🗼 🖻	B 10 8	AL ZI S	9 To 7 M	N *4	8 6 4 .	
E.	MS Sans Serif	• 10 •	BIU		2 - 4	- 2	- 1
Final Tum	or Histology				<u>.</u>		
							
Ifrom best of imac	jing. FNA. pathology. etc						
	,	· · · · · · · · · · · · · · · · · · ·					
Pathology	(if available)						
Tumor Size (cm):	by [
TMN Staging: T		. м	·	R:	F		
Earm View							
TMN Staging: T Form View	<u> </u>	<u> </u>		[R: [

Figure 45: Hepatocellular 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 46: Hepatocellular Cancer Pathology Table Schema

Hepatocellular Follow-up MR Follow-up MR Follow-up MR Follow-up MR Follow-up MR Follow-up MR Follow-up MR Follow-up Date of Visit: Weight (pounds): QOL score: ECOS performance status: © 0 1 2 © 3 © 4 U Veight (pounds): O © 1 © 2 © 3 © 4 © 2 © 3 © 4 © 0 © 1 © 2 © 3 © 4 © 2 © 3 © 4 • 2 © 3 © 4 • 2 © 3 © 4 • 2 © 3 © 4 • 2 © 4 • 4 • 4 • 4 • 4 • 5 • 6 • 7 • 6 • 7 • 7 • 7 • 7 • 8 • 1 • 8 • 1 • 9 • 1	File Edit View Insert Format	Records Tools Windo	ow Help Adobe PDF	Type a question for help	
Hepatocellular Selling Follow-up Information Add Record Find Record Delete Record MR Follow-up Window: Image: Comparison Image: Comparison Image: Comparison Date of Visit: Weight (pounds): QUL score: Image: Comparison Image: Comparison Date of Visit: Weight (pounds): QUL score: Image: Comparison Image: Comparison ECOG performance status: Image: Que Comparison Image: Que Comparison Image: Comparison Image: Comparison CEA: Albumin Alkaline phosphotase Image: Comparison Image: Comparison CEA: Albumin SGDT #Name? SGPT #Name? CEA: Albumin SGDT #Name? SGPT #Name? Redeveloped symptoms: Image: Comparison Image: Comparison Image: Comparison Weight loss Image: Comparison Image: Comparison Image: Comparison Image: Comparison Jaundice Vomiting Image: Comparison Image: Comparison Image: Comparison Image: Comparison Jaundice Vomiting Image: Comparison Image: Comparison Image: Comparison Image: Comparison <t< th=""><th>≝ • ⊌ •2 8 0 8 • ₩</th><th></th><th> 🍕 🖪 V 🕅 🕨 🕅</th><th> □ · □ · □ · □</th><th></th></t<>	≝ • ⊌ • 2 8 0 8 • ₩		🍕 🖪 V 🕅 🕨 🕅	□ · □ · □ · □	
Follow-up Information Add Record Find Record Delete Record MR Follow-up Window: Image: Construction of the second	Hepatocellular Spelling	er			
MR Follow-up Window: Date of Visit: Weight (pounds): ECDG performance status: 0 © 1 © 2 3 4 Lab Value: CEA: Albumin SGDT #Name? SGPT #Name? Redeveloped symptoms: Veright loss Biliary colic Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Clay colored stool Cholengitis Fatigue Status: O Died Death Date: O Net A.W.D., method of detection: Lab Radiologic evidence	Follow-up Informatio	n Add Record	Find Record Delete Reco	rd 🚺 🚺	
Min Date of Visit: Weight (pounds): QOL score: ECDG performance status: © 0 1 2 CEA: Albumin Alkaline phosphotase CA19-3: Total Bilirubin SGDT #Name? SGPT #Name? Redeveloped symptoms: Weight loss Biliary colic Pruritis Indigestion how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholangitis Fatigue Status: O Died Death Date: O Died Death Date: O AW.D., method of detection: Lab		Follow up Windows			
Date of Visit: Weight (pounds): QQL score: ECOG performance status: 0 1 2 3 4 Lab Value: CEA: Albumin Alkaline phosphotase CA19-3: Total Bilirubin SGOT #Name? Redeveloped symptoms: Weight loss Biliary colic how much (pounds) Nausea Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Specify: Cholangitis Fatigue Status: C Died Death Date: C N.E.D. C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence					
ECDG performance status: 0 1 2 2 3 4 Lab Value: CEA: Albumin CA19-3: Total Biliary colic Pruritis Indigestion how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Cholangitis Fatigue Status: C Died Death Date: C N.E.D. C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence	Date of Visit:	Weight (pounds):	QOL score:		
Lab Value: CEA: Albumin Alkaline phosphotase CA19-9: Total Bilirubin SGDT #Name? Redeveloped symptoms: SGDT #Name? SGPT #Name? Weight loss Biliary colic Pruritis Indigestion how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Specify: Cholangitis Fatigue Status: Colay colored stool Other Specify: Cholangitis Fatigue Colay colored stool Other Specify: Cholangitis Lab Radiologic evidence Clinical evidence	ECOG performance status: 🛛 🚳 0	@ 1 @ 2 @ 3	@ 4		
CEA: Albumin Alkaline phosphotase CA19-9: Total Bilirubin SGDT #Name? Redeveloped symptoms: SGPT #Name? Weight loss Biliary colic Pruritis how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Status: C Died Death Date: C N.E.D. C AW.D., method of detection: Lab Radiologic evidence Clinical evidence	Lab Value				
CA19-9: Total Bilirubin GA19-9: Total Bilirubin SGDT #Name? SGDT #Name? Redeveloped symptoms: Weight loss Biliary colic Pruritis Indigestion how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Status: Colice Died Death Date: Colice AW.D., method of detection: Lab Radiologic evidence Clinical evidence example Image: Colice Status: Colice Death Date: Colice AW.D., method of detection: Lab Radiologic evidence Clinical evidence			kalina phosphotasa		
CATION Substant Substant Redeveloped symptoms: Weight loss Biliary colic Pruritis how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Status: Cholengitis Fatigue O Died Death Date: Colecystidence C N.E.D. CAW.D., method of detection: Lab Radiologic evidence Clinical evidence		inubia 🗌 🤤		[#Namo2	
Redeveloped symptoms: Weight loss Biliary colic Pruritis Indigestion how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Cholangitis Fatigue		iiidbin j 3)		Hindlie:	
Weight loss Biliary colic Pruritis Indigestion how much (pounds) Nausea Abdominal pain Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Specify: Cholangitis Fatigue Status: Other Death Date: C N.E.D. N.E.D. C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence	Redeveloped symptoms:	15000-55055 DV	145-00-05 do 145-05	i i i i i i i i i i i i i i i i i i i	
how much (pounds) Jaundice Vomiting Back pain Cholecystitis Cholangitis Fatigue Status: O Died Death Date: C N.E.D. C A.W.D., method of detection: Radiologic evidence Clinical evidence	Veight loss	🔲 Biliary colic	🗖 Pruritis 🔲 🗌	ndigestion	
Jaundice Vomiting Back pain Cholecystitis Clay colored stool Other Specify: Cholangitis Fatigue Status: Other Death Date: C Died Death Date: C N.E.D. C A.W.D., method of detection: C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence	how much (pounds)	🗖 Nausea	🗖 Abdominal pain		
Cholecystitis Clay colored stool Other Specify: Cholangitis Fatigue Status: C Died Death Date: C N.E.D. C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence	🛛 🗖 Jaundice	🗖 Vomiting	🗖 Back pain		
Cholangitis Fatigue Status: C Died Death Date: C N.E.D. C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence	🗌 🗔 Cholecystitis	🗖 Clay colored stool	🗖 Other Specify:		
Status: O Died Death Date: O N.E.D. O A.W.D., method of detection: I Lab Radiologic evidence Clinical evidence	🗌 🗖 Cholangitis	📕 Fatigue			
C Died Death Date:	Status				
C N.E.D. C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence	C Died Death Date:			1	
C A.W.D., method of detection: Lab Radiologic evidence Clinical evidence					
	C AWD method of detections	Tush 🗖 Radiologia	avidance 🗖 Clinical avidanc		
scord: 14 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1					
erord: 14 4 1 1 1 1 1 1 1 1				<u>0</u> 2	
scord: 14 [4] 1] b [b [b #] of 1					
	ecord: 14 4 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
FUWin	Number	Follow-Up Windows
VisitDate	Date/Time	Visit Date
Weight	Number	Weight (lbs.)
QOLscore	Number	QoL Score (0-100)
ECOG	Number	ECOG Score (0-4)
LabCEA	Number	Laboratory - CEA
LabCA19-9	Number	Laboratory - CA19-9
LabAlb	Number	Laboratory - Albumin
LabBili	Number	Laboratory - Bilirubin
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
SxChola	Yes/No	Symptoms - Cholangitis
SxBC	Yes/No	Symptoms - Biliary Colic
SxNau	Yes/No	Symptoms - Nausea
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
SxAbd	Yes/No	Symptoms - Abdominal Pain
SxBack	Yes/No	Symptoms - Back Pain
SxOT	Yes/No	Symptoms - Other
SxOTSpe	Text	Symptoms - Other - Specify
Status	Number	Status (NED, AWD, Died)
DeathDate	Date/Time	Death Date
StatusAWDLab	Yes/No	AWD - Lab Evidence
StatusAWDRad	Yes/No	AWD - Radiology Evidence
StatusAWDCli	Yes/No	AWD - Clinical Évidence

Figure 48: Hepatocellular Cancer Follow-Up Table Schema

3.1.3 Gall Bladder/Biliary Cancer

	Unitat Records Tools			
L • 📙 😼 🎒 🗗 🖏 🖒	6 堕 唱 10 酱 2.	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1))* K @' ⊡ '@ ▼	2.
👻 MS Sans Serif	• 8 • B	<i>I</i> <u>U</u> ≡ ≡ ≡	2 · A · 2 · 🗌	• - • •
esumptive Diagno	sis at Onset of	Care		
coumpare blagne		<u></u>		
	<u> </u>			
esentation				
Late of Exaluation ECOG P	Performance Status	Height (in) Weig	nt (lbs)	
	1 @ 2 @ 3 @ 4		K (103.)	
ymptoms				
🗑 Weight Loss	🕅 Biliary Colic	🕅 Pruritis	Imdigestion	
How Much (pounds):	- 🕅 Nausea	🕅 Abdominal Pain	🕅 Dysphagia	
Jaundice	🕅 Vomiting	👿 Back Pain	🕅 Early Satiety	
≂ ol 1	🕅 Clay Colored Stool	🕅 Other Specify:		
Unolecystitis	🕅 Fatigue			
🖩 Cholecystitis 🗑 Cholangitis				
 Cholecystitis Cholangitis 				
 Cholecystitis Cholangitis 				
 Cholangitis 				
 Cholecystris Cholangitis 				

Figure 49: Gall Bladder/Biliary Cancer Presentation Form

Field Name	Data Type	Description
8D 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
5xSatiety	Yes/No	Initial Symptoms - Early Satiety
SxOT	Yes/No	Initial Symptoms - Other
SxOTSpe	Text	Initial Symptoms - Other - Specify
	100 800 Percent	

Figure 50: Gall Bladder/Biliary Cancer Presentation Table Schema

Lag File Ealt View Insert Pormat Records 10015 5	
Medical History	
Comorbidities Image: Heart Failure Image: Malnutrition Image: Ischemic Heart Disease Image: Liver Failure/Cirrhosis Image: Ischemic Heart Disease Image: Ischemic Heart Disease Image: Ischeart Di	Cancer History Patient Prior Dx: Image: Chemo Image: Chemo <t< td=""></t<>
Alcohol Use Envirormental Exposure Other - Specify:	Related Dx:

Figure 51: Gall Bladder/Biliary Cancer Medical History Form

ġ.	Field Name	Data Type	Description
81	(D	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
Ũ	CxHF	Yes/No	Comorbidities - Heart Failure
8	CXIHD	Yes/No	Comorbidities - Ischemic Heart Disease
	CxResp	Yes/No	Comorbidities - Respiratory
	CxDiab	Yes/No	Comorbidities - Diabetes
Ũ	CxDiabOral	Yes/No	Comorbidities - Diabetes - Insulin - Oral
8	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
1	CxHyper	Yes/No	Comorbidities - Hypertension
8	CxBleed	Yes/No	Comorbidities - Bleeding Disorder
	CxLiver	Yes/No	Comorbidities - Liver Failure
	CxMal	Yes/No	Comorbidities - Malnutrition
1	CxPriorCancer	Number	Comorbidities - Prior Cancer Dx
6	CxPriorCancerChemo	Yes/No	Comorbidities - Prior Cancer Dx - Chemo
	CxPriorCancerRadiation	Yes/No	Comorbidities - Prior Cancer Dx - Radiation
	CxPriorCancerSurgery	Yes/No	Comorbidities - Prior Cancer Dx - Surgery
1	SHCigarette	Yes/No	Social History - Cigarettes (significant use)
6	SHAlcohol	Yes/No	Social History - Alcohol (significant use)
	SHDrugUse	Yes/No	Social History - Drug Use
	SHExposure	Yes/No	Social History - Environmental Exposure
1	SHOther	Yes/No	Social History - Other
6	SHOtherS	Text	Social History - Other - Specify
	FamilyFatherDx	Number	Family History - Father Dx
	FamilyMotherDx	Number	Family History - Mother Dx
1	FamilyOther1	Text	Family History - Other1
8	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_5erum :	Form]				
🕄 File Edit	View Insert Format	<u>R</u> ecords <u>T</u> o	ols <u>W</u> indow <u>H</u> e	lp Ado <u>b</u> e PDF	Type a question fo	or help 🔽 🗕 🗗 🗙
🔛 - 🔛 🗞	🖨 🖪 🖤 🐰 🖻	B 0 8		5 7 4 >*	× 8 6 10	• 🛛 •
E .	MS Sans Serif	• 8 •	BIU	F = = 🆄	• 🔺 • 🖉 • 🔽	
Serum Stud	lies					
CEA: CA19-9:	Albumin Total Bilirubin	A A	ILK [IST: [ALT [Amylase: [
H. Pylori Status: C	Positive C Negative					
Form View						//

Figure 53: Gall Bladder/Biliary Cancer Serum Studies Form

	Field Name	Data Type	Description	
81	ID	AutoNumber	ID	
11°	MR	Text	Meditech Medical Record Number for Patient	
ŝ	LabCEA	Number	Laboratory - CEA	
	LabCA19-9	Number	Laboratory - CA19-9	
	LabAlb	Number	Laboratory - Albumin	
1	LabBili	Number	Laboratory - Bilirubin	
ŝ	LabAlka	Number	Laboratory - Alkaline phosphotase	
	LabALT	Number	Laboratory - ALT	
	LabAST	Number	Laboratory - AST	
1	LabAmylase	Number	Laboratory - Amylase	

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

Microsoft Access - [G]	8_3b_DiagIn	ng : Form]				
E File Edit View 1	nsert F <u>o</u> rma	at <u>R</u> ecords <u>j</u>	ools <u>W</u> indow	Help Add	be PDF Type a que	stion for help 🛛 🚽 🗗 🗙
	X B	000		97) (m • 2) .
					= 3 - A - 4	
Die an estie les suis					- 22 . 22 . 22	
Diagnostic imagii	ng Proce	aures				
Four Phase CT S	can or Ultr	asound Dur	<u>olex</u>			
Date of Procedure:	[
Status of Main PV:	l Open	Encased	@ Occluded	Abuts	Can't Tell	
Right Portal Vein:	l Open	Encased	C Occluded	Abuts	Can't Tell	
Left Portal Vein:	@ Open	@ Encased	@ Occluded	Abuts	🖉 Can't Tell	
Hepatic Arteries:	l Open	@ Encased	@ Occluded	Abuts	Can't Tell	
🕅 Mass Visible						
Extent of Involveme	nt of Hilar B	iliary Tree:				
🕅 Unknown	🕅 Right M	ain Duct 🖉 L	.eft Secondary Bi	anch 🕅	🛙 Left Tertiary Branch	
🕅 Left Main Duct	🕅 Bifurcati	on 🕅 F	Right Secondary I	Branch 🕅	🕅 Right Tertiary Branch	
Volumetric Assessment -	Proposed Live	er Remnant: 🕼	Done	oc @	Not Done	
Potential Maneuvers to	Increase Safet	y of Resection: @	Portal Vein Emt	olization @	Stents for Drainage @ N	lone
Chest X-Bay (CX	B)					0.000
E a						
Percutaneous I	ranshepat	<u>ic Cholangic</u>	ography (PT	<u>C1</u>		
Date of Procedure:	l Turner (© Jul	unal @ Eutom	i.			
i <u>ətenunu</u>	rype. v= ini	enidi ve Exterr				
Form View						

Figure 55: Gall Bladder/Biliary Cancer Diagnostic Imaging Form

Field Name	Data Type	Description
8) 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
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
	12	

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

Microsoft Access - [GB_4_Prelim : Form]				- O ×
E File Edit View Insert Format Records	<u>T</u> ools <u>W</u> indow	Help Adobe PDF	Type a question for help	×
₩· ₩ ₩ @ Q ♥ X № ® ∽		🦻 🔁 🕅 🕩	* 🕷 📴 🔂 😽 🖸).
	• B I U		2 • A • Z • 🔼 • [
Pre-Surgical Outlook				
Potentially Resectable	р. -			
Locally Advanced/Unresectable				
Metastatic or Equivocal Findings				
Form View				1

Figure 57: Gall Bladder/Biliary Cancer Preliminary Outlook Form

Field Nome	Data Type	Description
D 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)
2		

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

Picrosoft Access - [G	B_5_Treatment : Form]			
🗐 File Edit View	Insert Format Records	<u>I</u> ools <u>W</u> indow <u>H</u> elp Ad	lobe PDF Type a question fo	or help 🛛 🛨 🗕 🗗 🗙
🖳 • 🔚 🗞 🖨 🖸	1 🖤 👗 🖻 💼 🗠		🗚 🕨 🖉 🗗 🖕	• 🛛 •
*	*	• B I U 🗏 🗄	≡ <u></u> <u></u> · <u>A</u> · <u></u> .	· · · ·
Treatment Cou	irse			
Resection		Staging Laparos	copy/Laparotomy	
[™] <u>Radiation</u>		Palliative Measu	Ires	
Adjuvant Neoac	ljuvant 🖗 Both	🕅 Bypass	T HAL	
Chemotherapy		🕅 Gastrostomy Tube	🕅 PV Shunts	
l 🕼 Adjuvant 🖗 Neoad	ljuvant @ Both	💹 Jejunstomy Tube	🕅 Pall. Stenting	
🕅 Avastin	🕅 Leukovorin	🕅 Celiac Block	🕅 Pall. Radiation	
🕅 Capecitabine	🕅 Levamasole	Maracentesis	🕅 Pall. Resection	
🕅 Erbitux	Mitomycin	Thoracentesis	🕅 Other - Specify:	
🕅 Fluorouracil (5-FU)	🕅 Oxaliplatin	🕅 Transfusion		
FUDR	🕅 Taxol			
🕅 Gemcitabine	🕅 Other - Specify:	Experimental Pro	otocol (vaccine, etc.)	
🕅 Irinotecan		Genetic Counsel	ing	
	ta dh			
Form View				1

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]
😰 File Edit View Insert Format Records Tools Window Help Adobe PDF Type a question for help 🖌 🗗 🗙
K→BSBSVXBCSS824¥Ÿã▽A>>××®⊡a+Q.
- MS Sans Serif - 8 - B I U 画画画 公 - ム - ビー・
If Resection is Performed
Surgery Date of Admission: Procedure Type Image: Contract of Surgery: Image: Contract
IIII Venous Resection IIII Venous Reconstruction IIII Arterial Resection IIIII Arterial Reconstruction
Urner urgans nesected. Estimated Blood Loss (CC):
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 ELeak
Eiver Insufficiency (Total Bilirubin ≥ 5) If Yes. Total Bilirubin:
Date of Discharge: Discharge Status:
Form View

Figure 61: Gall Bladder/Biliary Cancer Resection Form

3	Field Name	Data Type	Description	
8	(D	AutoNumber	ID	
Uĩ	MR	Text	Meditech Medical Record Number for Patient	
1	ResDAdm	Date/Time	Resection - Date of Admission	
	ResDSurg	Date/Time	Resection - Date of Surgery	
<u>)</u> .	ResPxType	Number	Resection - Procedure Type (Whipple, total pancreatectomy, distal pancreatectomy, etc)	
11	ResORTime	Number	Resection - OR Time (hr.)	
1	ResVenRes	Yes/No	Resection - Venous Resection	
ŝ.	ResVenRec	Yes/No	Resection - Venous Reconstruction	
<u>)</u>	ResArtRes	Yes/No	Resection - Arterial Resection	
0	ResArtRec	Yes/No	Resection - Arterial Reconstruction	
25	ResOrgans	Text	Resection - Other Organs Resection	
á.	ResBloodLoss	Number	Resection - Estimated Blood Loss (cc)	
1	ResTransfusion	Yes/No	Resection - Tranfusion	
1	ResTUnits	Number	Resection - Transfusion Units	
8	ResTFFP	Yes/No	Resection - Transfusion - FFP	
ŝ.	ResTCell	Yes/No	Resection - Transfusion - Cell	
j.	ResAttempt	Number	Resection - Resection Attempt	
0	ResAttemptUn	Number	Resection - Resection Unsuccessful Reason (Tumor involvement, Operative mishap, etc)	
55	ResPOCourse	Number	Resection - PO - Post-Op Care Path	
<u>8</u>	ResPODays	Number	Resection - PO - Time in ICU (days)	
1	ResPOInfection	Yes/No	Resection - PO - Wound infection	
Û	ResPOLeak	Yes/No	Resection - PO - Leak	
8	ResPONG	Yes/No	Resection - PO - NG/gastrotomy drainage	
8	ResPOAbdominal	Yes/No	Resection - PO - Abdominal Collection	
]	ResPOPulmComp	Yes/No	Resection - PO - Pulminary Complications	
11	ResPOLiverInsuf	Yes/No	Resection - PO - Liver Insufficiency	
25-	ResPOLiverTB	Number	Resection - PO - Liver Insufficiency - Total Bilirubin	
6	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
2 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]	
E File Edit View Insert Format Records Iool	ols Window Help Adobe PDF Type a question for help 👻 🗕 🗗 🗙
₩ - B B B C V A B B P B	<u>≱</u> ;;, ÿ;;, ∀ M → ≪ 🕾 🗗 ′m • 🤉 .
MS Sans Serif 🔹 10 👻	₿ <i>Ⅰ</i> Щ ≣ ≣ ≣ Ѯ • <u>▲</u> • <u>⊿</u> • □ • .
Final Tumor Histology	
-	
(from best of imaging, FNA, pathology, etc)	
Pathology (if available)	
Tumor Size (cm): by	
TNM Staging: T 💽 N 💽 M	R: 💌
Form View	

Figure 65: Gall Bladder/Biliary Cancer Pathology Form

Field Name	Data Type	Description	
30 ID	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>R</u> ecords <u>T</u> ools <u>W</u> indow <u>H</u> elp Adobe PDF Type a question for help	8 ×
🔟 • 🔲 🖬 🖨 🖪 🖤 🐰	ⓑ í ⋈ 🍓 🛃 🏹 🏹 🚡 ▽ 🛤 > × 🗰 📾 + 🔍 🖕	
Gall Blagder/Biliary		
Follow-up Informatio	Add Record Find Record Delete Record	
MR #Name?	Follow-up Window:	
Date of Visit:	Weight (nounds)	
ECOG performance status:		
	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:		
C N.E.D.		
C A.W.D., Method of Detection:	🏾 Lab 🛛 🕅 Radiologic Evidence 🕅 Clinical Evidence	
1	22	
Form View		11.

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

Field Name	Data Type	Description
ID ID	AutoNumber	
MR	Text	Meditech Medical Record Number for Patient
FUWin	Number	Follow-Up Windows
VisitDate	Date/Time	Visit Date
Weight	Number	Weight (lbs.)
QOLscore	Number	QoL Score (0-100)
ECOG	Number	ECOG Score (0-4)
LabCEA	Number	Laboratory - CEA
LabCA19-9	Number	Laboratory - CA19-9
LabAlb	Number	Laboratory - Albumin
LabBili	Number	Laboratory - Bilirubin
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
SxChola	Yes/No	Symptoms - Cholangitis
SxBC	Yes/No	Symptoms - Biliary Colic
SxNau	Yes/No	Symptoms - Nausea
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
SxAbd	Yes/No	Symptoms - Abdominal Pain
SxBack	Yes/No	Symptoms - Back Pain
SxOT	Yes/No	Symptoms - Other
SxOTSpe	Text	Symptoms - Other - Specify
Status	Number	Status (NED, AWD, Died)
DeathDate	Date/Time	Death Date
StatusAWDLab	Yes/No	AWD - Lab Evidence
StatusAWDRad	Yes/No	AWD - Radiology Evidence
StatusAWDCli	Yes/No	AWD - Clinical Evidence

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

3.1.4 Gastric Cancer

Microsoft Access - [Gas_1_Prese	it : Form]				<u>-0×</u>
B <u>File E</u> dit <u>V</u> iew Insert Form	at <u>R</u> ecords <u>T</u> ools <u>Y</u>	<u>M</u> indow <u>H</u> elp Ado <u>b</u>	e PDF	Type a question for help	×
🖌 • 日 🕲 🍯 🖪 🖤 🐇 🛚	b 🛍 🗠 🍓 🎒	1 3 7 4	Ma Der MK 😭 d	🔁 🌆 🖌 😰 🗸	
MS Sans Serif	• 8 • B	/ <u>U</u> ≣≣≣	I <u>⊉ • </u> ▲ • _	<u>/</u>	
resumptive Diagnosi	s at Onset of (Care			
	-1				
Presentation					
Date of Evaluation ECOG Perfo	mance Status	<u>Height (in.) Wei</u>	aht (lbs.)		
1	10 2 10 3 10 4				
Symptoms	-				
100 Historick Loca	I Biliary Colic	I Pruntis	III Indigestion		
Im Weight Loss	0.002 20.00	00002 W 25 25 020 25	Charles of the second sec		
How Much (pounds):	🕅 Nausea	🕅 Abdominal Pain	🕅 Dysphagia		
How Much (pounds):	I≣ Nausea I≣ Vomiting	👿 Abdominal Pain 👿 Back Pain	👿 Dysphagia 🕅 Early Satiety		
How Much (pounds):	 Nausea Vomiting Clay Colored Stool 	 Abdominal Pain Back Pain Other Specify: 	👿 Dysphagia 🕅 Early Satiety		
How Much (pounds): Jaundice	 Nausea Vomiting Clay Colored Stool Fatigue 	III Abdominal Pain IIII Back Pain IIII Other Specify:	I Dysphagia I Early Satiety		

Figure 69: Gastric Cancer Presentation Form

3	Field Name	Data Type	Description
8	ID	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
1	PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)
8 3	DemEvalDate	Date/Time	Demographics - Date Evaluated by Surgical Oncology
<u>.</u>	DemECOG	Number	Demographics - ECOG Score (0-4)
[]	DemHeight	Number	Demographics - Height in Inches of Patient
54 	DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission
8	SxWtloss	Yes/No	Initial Symptoms - Weight Loss
1.	SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds
11	SxJaun	Yes/No	Initial Symptoms - Juandice
1	SxChole	Yes/No	Initial Symptoms - Cholecystitis
\$ 2	SxChola	Yes/No	Initial Symptoms - Cholangitis
	SxBC	Yes/No	Initial Symptoms - Biliary Colic
	SxNau	Yes/No	Initial Symptoms - Nausea
0.6%	SxVom	Yes/No	Initial Symptoms - Vomiting
8	SXCCS	Yes/No	Initial Symptoms - Clay Colored Stool
<u>.</u>	SxFati	Yes/No	Initial Symptoms - Fatigue
11	SxPru	Yes/No	Initial Symptoms - Pruritis
1	SxInd	Yes/No	Initial Symptoms - Indigestion
8	SxAbd	Yes/No	Initial Symptoms - Abdominal Pain
Į.,	SxBack	Yes/No	Initial Symptoms - Back Pain
11	SxDyspha	Yes/No	Initial Symptoms - Dysphagia
2	SxSatiety	Yes/No	Initial Symptoms - Early Satiety
8	SxOT	Yes/No	Initial Symptoms - Other
4.	SxOTSpe	Text	Initial Symptoms - Other - Specify

Figure 70: Gastric Cancer Presentation Table Schema

Medical History				• □••	
Comorbidities Image: End of the en	Malnutrition Liver Failure/Cirrhosis Less than Six Months Greater than Six Months Greater than Six Months Dial Agents Diet Control trug Use ntal Exposure	Cancer History Patient Prior Dx: Chemo Chemo Father Dx: Cher Relation: Related Dx: Cher Relation: Related Dx: Related Dx: Related Dx: Cher Relaton: Related Dx: Cher Relaton: Related Dx: Cher R	on T Surgery		

Figure 71: Gastric Cancer Medical History Form

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

Figure 72: Gastric Cancer Medical History Table Schema

🖉 Microsoft A	Access - [Gas_3a_Serum	: Form]					_ _ ×
🗄 <u>File E</u> di	t <u>V</u> iew Insert F <u>o</u> rmat	<u>R</u> ecords <u>I</u>	ools <u>W</u> indo	w <u>H</u> elp Ado <u>b</u> e	PDF	Type a question for help	8 ×
🛛 🕶 🖬 🖗	8 6 R 🖤 🕺 B	B 10 8		y to v	• >* W 😭	🗗 🔚 📲 🖳 🚬	
	MS Sans Serif	• 8	BZ	U E E E	2 - A -	4 · · · · · ·	
Serum St	udies		-				
CEA:	Albumin		ALK	ALT			
CA19-9:	Total Bilirubin		AST:	Amyl	lase:		
H. Pylori Status	C Positive C Negative						
Form View							1 1 2

Figure 73: Gastric Cancer Serum Studies Form

Field Name	Data Type	Description
8) 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 74: Gastric Cancer Serum Studies Table Schema

Microsoft Access - [Gas_3b_DiagImg :	Form]		- O ×
🗐 File Edit View Insert Format F	ecords <u>T</u> ools <u>W</u> indow <u>H</u> elp Ado <u>b</u> e PDF	Type a question for help	8 ×
1×- = = = = = = = = = = = = = = = = = =		∎ /m • 12.	
		2	
Diagnostic Imaging Procedur	25		
CT with Panreatic Protocol/CT/	× ·		
Date of Procedure:			
Celiac Artery Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknow	in	
Superior Mesenteric Artery Involvement	19 Open 19 Abutted 19 Encased 19 Occluded 19 Unknow	in	
🕅 Hepatic Artery Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknow	'n	
📓 Inferior Vena Cava Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknow	m	
🕅 Superior Mesenteric Vein Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknow	m	
🕅 Portal Vein Involvement	@ Open @ Abutted @ Encased @ Occluded @ Unknow	m	
Nodes: 🖉 Celiac Nodal Disease	🖉 Other Nodal Disease 🖉 No Nodal Assessment or Mentio	n	
Tumor Size (cm): by			
Chest X-Bay (CXB)			
Percutaneous Transhepatic Ch	olangiography (PTC)		
Date of Procedure:	-		
IIIII <u>Stenting</u> Type: ^{(%} Internal	@ External		
Form View			/

Figure 75: Gastric Cancer Diagnostic Imaging Form

Field Name	Data Type	Description
8 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

5	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

	\♥ X��� × 	○ ⑧ 월 월 월 ♡ ▶ ▷ ₩ ▶ ₩ ☎ ◘ 面 ♡. ₽ ℤ 및 ≣ ≣ ≣ ⊉ · ▲ · ℤ · □· □·.
Treatment Cou	<u>irse</u>	
	juvant @ Both III Leukovorin III Leukovorin III Levamasole III Mitomycin III Oxaliplatin	Staging Laparoscopy/Laparotomy Palliative Measures Bypass HAL Gastrostomy Tube PV Shunts Jejunstomy Tube Pall. Stenting Celiac Block Pall. Radiation Paracentesis Pall. Resection Thoracentesis Other - Specify: Transfusion Tensfusion
FUDR Gemcitabine Till Irinotecan	I Taxol I Other - Specify:	Experimental Protocol (vaccine, etc.) Genetic Counseling

Figure 79: Gastric Cancer Treatment Form

Field Name	Data Type	Description
₿▶ ID	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 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	×
₩ - ₩ ∰ ∰ & ™ & ™ ® 0 % A N ≫ T > T # > × × ® @ - Q .	
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 Image: Successful Generation Post-Op Image: Successful - Reason: Days in ICU: Image: Successful - Reason: Post-Op Care Path: Image: Congruent Image: Divergent	
🕅 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: Discharge Status:	
	-
Form View	

Figure 81: Gastric Cancer Resection Form

3	Field Name	Data Type	Description	
8	D	AutoNumber	ID	
Ű	MR	Text	Meditech Medical Record Number for Patient	
1	ResDAdm	Date/Time	Resection - Date of Admission	
	ResDSurg	Date/Time	Resection - Date of Surgery	
<u>)</u> .	ResPxType	Number	Resection - Procedure Type (Whipple, total pancreatectomy, distal pancreatectomy, etc)	
11	ResORTime	Number	Resection - OR Time (hr.)	
1	ResVenRes	Yes/No	Resection - Venous Resection	
ŝ.	ResVenRec	Yes/No	Resection - Venous Reconstruction	
<u>.</u>	ResArtRes	Yes/No	Resection - Arterial Resection	
1	ResArtRec	Yes/No	Resection - Arterial Reconstruction	
25	ResOrgans	Text	Resection - Other Organs Resection	
á.	ResBloodLoss	Number	Resection - Estimated Blood Loss (cc)	
1	ResTransfusion	Yes/No	Resection - Tranfusion	
11	ResTUnits	Number	Resection - Transfusion Units	
1	ResTFFP	Yes/No	Resection - Transfusion - FFP	
ŝ.	ResTCell	Yes/No	Resection - Transfusion - Cell	
<u>.</u>	ResAttempt	Number	Resection - Resection Attempt	
0	ResAttemptUn	Number	Resection - Resection Unsuccessful Reason (Tumor involvement, Operative mishap, etc)	
8.C	ResPOCourse	Number	Resection - PO - Post-Op Care Path	
13. 13.	ResPODays	Number	Resection - PO - Time in ICU (days)	
1	ResPOInfection	Yes/No	Resection - PO - Wound infection	
11	ResPOLeak	Yes/No	Resection - PO - Leak	
8	ResPONG	Yes/No	Resection - PO - NG/gastrotomy drainage	
ģ.	ResPOAbdominal	Yes/No	Resection - PO - Abdominal Collection	
<u>.</u>	ResPOPulmComp	Yes/No	Resection - PO - Pulminary Complications	
1	ResPOLiverInsuf	Yes/No	Resection - PO - Liver Insufficiency	
5.C	ResPOLiverTB	Number	Resection - PO - Liver Insufficiency - Total Bilirubin	
63 33	ResPODDischarge	Date/Time	Resection - Date of Discharge	
1	ResPODischStatus	Number	Resection - Discharge Status	

Figure 82: Gastric Cancer Resection Table Schema

Microsoft Access - [Gas_6b_NoRes : Form]			_ 🗆 🗙
E File Edit View Insert Format Records	<u>I</u> ools <u>W</u> indow <u>H</u> elp Ado <u>b</u> e Pl	DF Type a question for help	×
🔟 - 🖬 🖏 🎒 🖪 🖤 👗 🖻 💼 🕫)* X 🗗 🗗 🗐 🔹 🕄 .	
*	• B I U E E E	<u>2</u> · <u>A</u> · <u>I</u> · <u>-</u> ·	
If Resection is Not Performed			
Date of Decision:			
Reasons (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			
Vascular Involvement	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			
Form View			

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 - [Gas_7_Path : Form]	-OX
Ele Edit Yiew Insert Format Records Iools Window Help Adobe PDF Type a question for help	8 ×
IM·日电 每 Q、♥ % № C い % A X V T V T M H* M P D 1 · Q.	
MS Sans Serif • 10 • B I U ≣ ≣ ≣ 2 • ▲ • 2 • □ • •	
Final Tumor Histology	
(from best of imaging, FNA, pathology, etc)	
Pathology (if available)	
Tumor Size (cm): by	
TNM Staging: T N M R:	
Form View	/_

Figure 85: Gastric 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	tion - 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
Effe Edit Yiew Insert Format Records Iools Window Help Adobe PDF Type a question for help _ 5 W Image: Second Se	×
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:	
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 Voniting 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: I Lab III Radiologic Evidence III Clinical Evidence	
Form View	11

Figure 87: Gastric Cancer Follow-Up Form

Field Name	Data Type	Description	
ID	AutoNumber		
MR	Text	Meditech Medical Record Number for Patient	
FUWin	Number	Follow-Up Windows	
VisitDate	Date/Time	Visit Date	
Weight	Number	Weight (lbs.)	
QOLscore	Number	QoL Score (0-100)	
ECOG	Number	ECOG Score (0-4)	
LabCEA	Number	Laboratory - CEA	
LabCA19-9	Number	Laboratory - CA19-9	
LabAlb	Number	Laboratory - Albumin	
LabBili	Number	Laboratory - Bilirubin	
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	
SxChola	Yes/No	Symptoms - Cholangitis	
SxBC	Yes/No	Symptoms - Biliary Colic	
SxNau	Yes/No	Symptoms - Nausea	
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	
SxAbd	Yes/No	Symptoms - Abdominal Pain	
SxBack	Yes/No	Symptoms - Back Pain	
SxOT	Yes/No	ymptoms - Other	
SxOTSpe	Text	Symptoms - Other - Specify	
Status	Number	Status (NED, AWD, Died)	
DeathDate	Date/Time	Death Date	
StatusAWDLab	Yes/No	AWD - Lab Evidence	
StatusAWDRad	Yes/No	AWD - Radiology Evidence	
StatusAWDCli	Yes/No	AWD - Clinical Evidence	

Figure 88: Gastric Cancer Follow-Up Table Schema

3.1.5 Esophageal Cancer

E File Edit :		esent : I	orm]												-10
NA THERE	View Insert	Format	Records	Tools	Windo	w <u>H</u> e	lp Ad	o <u>b</u> e PDF		Туре	a qu	Jestia	on for he	elp.	8
🗮 🖌 📕 🖾	🖨 🖪 🖤 🛛	X 电		. 2	↓ Z↓	V Y	i V	4	* 195		E	5 1/1	- 2	9.	
÷	MS Sans Serif		• 8	• B	I	<u>u</u>	F 🐺	≡ ≤	2 - 1	<u>+</u> -	1	• [-	- •	-
Presumptiv	e Diagn	osis a	at Ons	et of	Car	е									
1000	(1999) (1999)		Ŧ	ſ											
Presentati	on														
Date of Evalua	ation ECOG	Perform	ance Stat	us	Heig	<u>ht (in.)</u>	W	eight (l	<u>bs.</u>]						
	60	© 1 ©	2 @ 3	@ 4	1	1	L								
Symptoms		-						-							
I Weight Loss		_ 🕅	Biliary Col	ic	[***	Pruritis			Indige	estion					
How Much (p	oounds):		Nausea		I	Abdom	inal Pair	n 🔎	i Dyspł	nagia					
🕅 Jaundice			Vomiting		I	Back F	ain	٦.	Early	Satiety					
		l.	Clay Color	red Stool	I.	Other	Specif	y:							
Cholecystitis		-	Entimum			1	- <u></u>	-			16				

Figure 89: Esophageal Cancer Presentation Form

3	Field Name	Data Type	Description
8	ID	AutoNumber	ID
	MR	Text	Meditech Medical Record Number for Patient
	PresumptiveDx	Number	Presumptive Diagnosis (Pancreatic tumor, periampullary tumor, etc)
8 -	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
6 I	SxWtloss	Yes/No	Initial Symptoms - Weight Loss
	SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds
11	SxJaun	Yes/No	Initial Symptoms - Juandice
	SxChole	Yes/No	Initial Symptoms - Cholecystitis
8 -	SxChola	Yes/No	Initial Symptoms - Cholangitis
	SxBC	Yes/No	Initial Symptoms - Biliary Colic
	SxNau	Yes/No	Initial Symptoms - Nausea
1000	SxVom	Yes/No	Initial Symptoms - Vomiting
8	SXCCS	Yes/No	Initial Symptoms - Clay Colored Stool
	SxFati	Yes/No	Initial Symptoms - Fatigue
11	SxPru	Yes/No	Initial Symptoms - Pruritis
	SxInd	Yes/No	Initial Symptoms - Indigestion
8 I	SxAbd	Yes/No	Initial Symptoms - Abdominal Pain
	SxBack	Yes/No	Initial Symptoms - Back Pain
1	SxDyspha	Yes/No	Initial Symptoms - Dysphagia
Сî	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

Elle Edit Yew Insert Figmat Records Iools Window Help Adobe PDF Type a question for help -	Microsoft Access - [Eso_2_History : Form]	×
Comorbidities Cancer History Image: Heart Failure Image: Malnutrition Patient Prior Dx: Image: Ischemic Heart Disease Interpretation Image: Interpretation Image: Ischemic Heart Disease Image: Interpreta	E File Edit View Insert Format Records Iools	Window Help Adobe PDF Type a question for help _ # # X Y </th
IIII Other - Specify: Related Dx:	Comorbidities Image: Heart Failure Image: Malnutrition Image: Science Heart Disease Image: Liver Failure/Cirrhosis Image: Science Heart Dis	Cancer History Patient Prior Dx: Im Chemo

Figure 91: Esophageal Cancer Medical History Form

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

Figure 92: Esophageal Cancer Medical History Table Schema

PMicrosoft Acc	ess - [Eso_3a	_Serum : I	Form]									
EB Eile Edit	⊻iew Insert	Format	Records	Tools	Window	Help	Ado <u>b</u> e P	DF	Туре а	question fo	or help	×
🔟 • 🔛 🗞	a 🕽 💞	光 暭			↓ Z↓ ĭ	y 🚡	7 4)* MX	P	6 'a •	2.	
-	MS Sans Ser	if	• 8	• B			≣ ≇	2 - 1	<u>.</u> .	/ • [- 1
Serum Stud	ies									_		
CEA:	Albumi	n [ALK			ALT					
CA19-9:	Total E	lilirubin 🛛		AST:			Amylas	se:	1			
H. Pylori Status: 🤿	Positive C	Negative										
Form View												1.

Figure 93: Esophageal Cancer Serum Studies Form

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

Figure 94: Esophageal Cancer Serum Studies Table Schema

Microsoft Access - [Eso_3b_DiagImg :	Form]	- 🗆 ×
E Edit View Insert Format	<u>ecords Tools Window Help Adobe PDF</u> Type a question for help 👻	_ 8 ×
	a ∽ @ # # ¥ ™ a ∨ m ↦ ₩ a ⊡ a • Q . - B # ¥ ≡ ≡ ■ 2 • A • 2 • □ • □ • .	
Diagnostic Imaging Procedur	es	
CT with Panreatic Protocol/CT/		
Date of Procedure:		
Celiac Artery Involvement	C Open C Abutted C Encased C Occluded C Unknown	
Superior Mesenteric Artery Involvement	© Open @ Abutted @ Encased @ Occluded @ 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	C Open C Abutted C Encased C Occluded C 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 Date of Procedure: Im Stenting Type: @ Internal	eolangiography (PTC) © External	
Form View		

Figure 95: Esophageal Cancer Diagnostic Imaging Form

Field Name	Data Type	Description
B 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

ġ.	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 98: Esophageal Cancer Preliminary Outlook Table Schema

Resection Staging Laparoscopy/Laparotomy Radiation Palliative Measures	Treatment Co	<u>urse</u>		<u>= ~ . ~ . ~ .</u>	
Radiation Palliative Measures [©] Adjuvant [©] Neoadjuvant [©] Both [©] Bypass [©] HAL [©] Adjuvant [©] Neoadjuvant [©] Both [©] Gastrostomy Tube [©] PV Shunts [©] Adjuvant [©] Neoadjuvant [©] Both [©] Gastrostomy Tube [©] Pall Stenting [©] Adjuvant [©] Neoadjuvant [©] Both [©] Celiac Block [©] Pall Radiation [©] Capecitabine [©] Levamasole [©] Paracentesis [©] Pall Resection [©] Etbitux [©] Mitomycin [©] Thoracentesis [©] Other - Specify: [©] FUDR [©] Taxol [©] Transfusion [©] Transfusion	E <u>Resection</u>		Staging Laparos	copy/Laparotomy	
Chemotherapy If Gastrostomy Tube If PV Shunts Adjuvant @ Neoadjuvant @ Both If Gastrostomy Tube If PV Shunts Avastin If Leukovorin If Celiac Block If Pall. Stenting Capecitabine If Leukovorin If Celiac Block If Pall. Resection If Erbitux If Mitomycin If Thoracentesis If Other - Specify: If FUDR If Taxol If Taxol	Radiation Adjuvant Neoa	djuvant 🖗 Both	Palliative Measu	Ires	
Image: Second	Chemotherapy	duuset @ Dette	Gastrostomy Tube	PV Shunts	
	Mujuvani ve Neba	E Leukovorin	IIII Jejunstomy Lube IIII Celiac Block	I≋ Pall. Stenting I≅ Pall. Radiation	
IIII Fluorouracii (5+FU) IIII Uxaliplatin IIII Transfusion IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	Erbitux		IIII Paracentesis IIII Thoracentesis	IIII Pall. Resection	
Experimental Protocol (vaccine etc.)	Fluorouracii (5+0)	I≋ Uxaliplatin I≅ Taxol	Fransfusion	tocol (vaccine, etc.)	
Image: Second	I≋ Gemcitabine I≣ Irinotecan	III Other - Specify:	Genetic Counsel	ing	

Figure 99: Esophageal Cancer Treatment Form

Field Name	Data Type	Description
₿▶ ID	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 - Fluorouracii (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 100: Esophageal Cancer Treatment Table Schema

Microsoft Access - [Eso_6a_Res : Form]
Image: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF Type a question for help • _ 6 > Image: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF Image: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF Image: Mage: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF Image: Mage: Serie Format Records Tools Window Help Adobe PDF Mage: Serie Format Records Tools Window Help Adobe PDF <t< th=""></t<>
If Resection is Performed
Date of Admission: Date of Surgery: Procedure Type Image: Construction of the construction
Resection Attempt: © Successful © Unsuccessful - Reason: Post-Op Days in ICU: Post-Op Care Path: © Congruent © Divergent
Im NG/Gastrostomy Drainage > 7days Im Abdominal Collection Im Prolonged Ileus Im Pulmonary Complications Im Wound Infection Im Small Bowel Obstruction
Image: Control Bilinubin > 5) If Yes, Total Bilinubin: Date of Discharge: Discharge Status:
Form View

Figure 101: Esophageal Cancer Resection Form

Field Name	Data Type	Description
₿▶1D	AutoNumber	ID and an an an and
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	
▶)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 106: Esophageal Cancer Pathology Table Schema

🔎 Microsoft Access - [Follow-up In	formation]			
📅 Eile Edit Yiew Insert For	mat <u>R</u> ecords <u>T</u> ools <u>W</u> indo	w <u>H</u> elp Ado <u>b</u> e PDF	Type a question for help	×
		V 1 V 1		
Esophageal Tumor				
Follow-up Informatio	n Add Record Fin	Record Delete Record	N •	
MB 000726070	Follow-up Window:	9M •		
Date of Visit 05/11/2003	Weight (pounds):	QOL score:		
ECOG performance status:	@1 @2 @3 @	4		
Lah Value:				
CEA: Albumin	Alkalir	e Phosphotase		
CA19-9: Total Bill	irubin	aon construction		
Redeveloped Symptoms:				
🗖 Weight loss	🗖 Biliary colic	🗖 Pruritis 🛛 🗖 Bac	k pain	
how much (pounds)	Nausea	Abdominal pain 🗖 Ind	igestion	
L Jaundice	Clau colored stool	Uther Specity:		
Cholangitis	Fatigue	1		
Challen				
C Diad Dath Data	E /11/2002			
	0711/2003			
C A.W.D., Method of Detection:	🛛 Lab 🛛 🗖 Radiologic Evid	ence 🔲 Clinical Evidence		
22				
Form View				11.

Figure 107: Esophageal Cancer Follow-Up Form

Field Name	Data Type	Description
8 ID	AutoNumber	
MR	Text	Meditech Medical Record Number for Patient
FUWin	Number	Follow-Up Windows
VisitDate	Date/Time	Visit Date
Weight	Number	Weight (bs.)
QOLscore	Number	QoL Score (0-100)
ECOG	Number	ECOG Score (0-4)
LabCEA	Number	Laboratory - CEA
LabCA19-9	Number	Laboratory - CA19-9
LabAlb	Number	Laboratory - Albumin
LabBili	Number	Laboratory - Bilirubin
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
SxChola	Yes/No	Symptoms - Cholangitis
SxBC	Yes/No	Symptoms - Biliary Colic
SxNau	Yes/No	Symptoms - Nausea
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
SxAbd	Yes/No	Symptoms - Abdominal Pain
SxBack	Yes/No	Symptoms - Back Pain
SxOT	Yes/No	Symptoms - Other
SxOTSpe	Text	Symptoms - Other - Specify
Status	Number	Status (NED, AWD, Died)
DeathDate	Date/Time	Death Date
StatusAWDLab	Yes/No	AWD - Lab Evidence
StatusAWDRad	Yes/No	AWD - Radiology Evidence
StatusAWDCli	Yes/No	AWD - Clinical Evidence
	10.004000	

Figure 108: Esophageal Cancer Follow-Up Table Schema

3.1.6 Colorectal Cancer

• 🖬 🖬 📾 🕼 🗸 🗠	8 · 10 · 16 10 % Ž↓ X↓ У 13 ∨ #1 ▶* ** 11 11	
*	→ B I U 臣 吾 王 公・A・ Z・	
esumptive Diagno	osis at Onset of Care	
Sino D		
esentation		
		1
ate of Evaluation ECUG F	<u>Performance Status Height [in.] Weight [ibs.]</u>	
umotome		
Veight Loss	🖩 Bloating 🖉 BBBPB 🖉 Constination	
How Much (pounds):	Im Distance Im Distan	
	🗑 Vomiting 🖉 Back Pain 🕅 Diarrhea	
🗄 Anal Pain		
CAnal Pain CPerineal Pain	188 Bowel Habit Change 188 Uther Specify:	
≋ Anal Pain ፪ Perineal Pain ፪ Bowel Obstruction	IIII Bowel Habit Change IIII Uther Specify:	

Figure 109: Colorectal Cancer Presentation Form

0	Field Name	Data Type	Description	8
81	(D	AutoNumber	ID	
	MR	Text	Meditech Medical Record Number for Patient	
Î.	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	
Į.	SxWtlossP	Number	Initial Symptoms - Weight Loss - Pounds	
	SxAnalPain	Yes/No	Initial Symptoms - Anal Pain	
Î.	SxPerPain	Yes/No	Initial Symptoms - Perineal Pain	
	SxBObs	Yes/No	Initial Symptoms - Bowel Obstruction	
	SxBloat	Yes/No	Initial Symptoms - Bloating	
	SxNau	Yes/No	Initial Symptoms - Nausea	
i –	SxVom	Yes/No	Initial Symptoms - Vomiting	
	SxBHabit	Yes/No	Initial Symptoms - Bowel Habit Change	
	SxFati	Yes/No	Initial Symptoms - Fatigue	
	SxBRBPR	Yes/No	Initial Symptoms - BRBPR	
	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	
Ú.	SxDiar	Yes/No	Initial Symptoms - Diarrhea	
	SXOT	Yes/No	Initial Symptoms - Other	
1	SxOTSpe	Text	Initial Symptoms - Other - Specify	

Figure 110: Colorectal Cancer Presentation Table Schema

Microsoft Access - [Col_2_History : Form]	
Elle Edit View Insert Format Records Io ▲ • ■ • ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●	ols Window Help Adobe PDF _ &
Comorbidities Im Heart Failure Im Ischemic Heart Disease Im Ischemic Heart Disease Im Ischemic Heart Disease Im Diabetes Im Perspiratory Im Diabetes Im Perspiratory Im Diabetes Im Perspiratory Im Diabetes Im Diabetes Im Perspiratory Im Diabetes Im	Cancer History Patient Prior Dx Patient Prior Dx Image: Chemo Image: Chemo Father Dx: Image: Chemo Mother Dx: Other Related Dx: Other Related Dx: Image: Chemo Related Dx:
Form View	

Figure 111: Colorectal Cancer Medical History Form

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

Figure 112: Colorectal Cancer Medical History Table Schema

Microsoft A	ccess - [Col_3a_Serum	Form]				-OX
🖪 Eile Edit	<u>V</u> iew Insert Format	Records Ioc	ols <u>W</u> indow	Help Adobe PDF	=	_ 8 ×
🔟 • 🔲 🤋	3 🗿 🖪 🖤 🕺 Pe	B 0 8		× 7 4	• * * 🖻 🗗 'a •	2.
line .	- MS Sans Serif	• 8 •	BIU		≥ • ▲ • ⊿ • □•	···
Serum St	tudies					
CEA:	Albumin	A	lk 「	ALT		
CA19-9:	Total Bilirubin	A	ST:	Amylase:		
Form View						1

Figure 113: Colorectal Cancer Serum Studies Form

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

Figure 114: Colorectal Cancer Serum Studies Table Schema

Microsoft Access - [Col_3b_Diag : Fo	rm]	
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):	
		<u> </u>
Form View		

Figure 115: Colorectal Cancer Diagnostic Imaging Form

Field Name	Data Type	Description	
8 ID	AutoNumber	ID	
MR	Text	Meditech Medical Record Number for Patient	
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	
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: C Percutaneous RFA Reasons
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 Pesection and/or Ablation Abandoned Variation of Planned Resection and/or Ablation Pump Added Too Close, or Positive Margin Resection Added to Operation More Resection Done Resection Added to Operation More Resection Done Pump Not Done
Form View

Figure 117: Colorectal Cancer Preliminary Outlook Form

	Save Jame	Data Type	Description
8		AutoNumber	ID
100	MR	Text	Meditech Medical Record Number for Patient
1	PreOutlook	Number	Pre-Surgical Tumor Outlook (Potentially Resectable, Locally Advanced/Unresectable, Metastatic
	PET	Yes/No	PET Scan
	SpecialCT	Yes/No	Special CT
	Percut	Number	Percutaneous Ablation -
î.	PercutN	Number	Percutaneous Ablation - Should Not - Reasons
	PercutY	Number	Percutaneous Ablation - Could Be - How
	PercutYRFA	Number	Percutaneous Ablation - Could Be - RFA
	Techniq	Number	Specialized Techniques
Û.	Lap	Number	Laparoscopic Operation Planned
	Operation	Number	Other Operation Planned
	Plan	Yes/No	Standard Operations Performed
	PlanY	Number	Standard Operations Performed - Types
1	PlanYR	Number	Standard Operations Performed - Resection
	PlanNRes	Number	Standard Operations Not Performed - Reason
	PlanNType	Number	Standard Operations Not Performed - Result
	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
	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
	PlanNType_PND	Yes/No	Standard Operations Not Performed - Result - Pump Not Done

Figure 118: Colorectal Cancer Preliminary Outlook Table Schema

Microsoft Access - [C]	ol_5_Treatment : Form]			
E Eile Edit View	Insert Format Records	Iools Window Help Ad	lobe PDF Type a question for help	• _ 8 ×
× - B 8 6 0	à.♥ X @ @ ~		M >* W 🗗 🗗 🗤 🖸).
*	*	• B I U = =	≡ <u>></u> · <u>∧</u> · <u>√</u> · <u></u> ·	- • •
Treatment Course				
Resection		Staging Laparos	copy/Laparotomy	
Radiation		🗏 Palliative Measu	ires	
l 🖉 Adjuvant 🖗 Neoadjuvant 🖗 Both		Bypass	The 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	🖉 Europianostal Der		
🕅 Gemcitabine	🕅 Other - Specify:			
🕅 Irinotecan	Genetic Counseling			
Form View			ni niv ni niv ni niv	11.

Figure 119: Colorectal Cancer Treatment Form
Field Name	Data Type	Description					
₿▶ ID	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	Freatment - 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					
TxPalO5	Text	Treatment - Palliation - Other - Specify					
TxExp	Yes/No	Treatment - Experimental protocol (ie. vaccine)					
TxGene	Yes/No	Treatment - Gene Counseling					

Figure 120: Colorectal Cancer Treatment Table Schema





1	Field Name	Data Type	Description	
81	ID	AutoNumber	ID	
10	MR	Text	Meditech Medical Record Number for Patient	
1	Per_DIH	Number	Days in Hospital	
	Per_ICU	Number	Days in ICU	
	Per_Trans	Yes/No	Blood Transfusion	
10	Per_Trans_Units	Number	Units	
1	Per_chest	Yes/No	Chest Tube	
1	Per_pneu	Yes/No	Pneumothorax	
	Per_emb	Yes/No	Embolization	
10	Per_sub	Yes/No	Subsequent Drainage	
Ť.	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 🗸 🗗 🤉
😫 • 🖬 ₦ 🖨 Q, ♥ ¾ ħ ħ ħ ⋈ ⋈ 🍕 🏭 Ў ħ ♡ 🛤 > × ⋈ 🗗 🗗 Ѣ • Q,
MS Sans Serif ・ 8 ・ B I U 臣 吾 吾 ②・ ▲・ 2・ □・ .
If Resection is Performed
Date of Admission: Date of Surgery: OR: Time (hr):
Procedure Type Days in ICU:
🕼 CVP cm/H2D During Hepatic Transection/Puncture: cm/h2D:
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

5	Field Name	Data Type	Description	8
81)D	AutoNumber	ID	
	MR	Text	Meditech Medical Record Number for Patient	
1	DAdm	Date/Time	Date of Admission	
3	DSurg	Date/Time	Date of Surgery	
	ORTime	Number	OR Time	
	PxType	Number	Procedure Type	
1	Organs	Text	Other Organs Resected	
8	BloodLoss	Number	Blood Loss (cc)	
	CVP	Yes/No	CVP cm/H20 During Hepatic Transection/Puncture	
	CVPcm	Number	CVP cm/H20	
1	Tidalcc	Number	Tidal Volume During Hepatic Transection/Puncture (cc)	
3	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	
8	T_Tissue	Yes/No	Transfusion - Tissue Link	
	T_CVSA	Yes/No	Transfusion - CVSA	
	T_Finger	Yes/No	Transfusion - Finger Fracture	
1	T_Clamp	Yes/No	Transfusion - Clamp/Crush	
8	T_Argon	Yes/No	Transfusion - Argon Beam	
	T_Pary	Yes/No	Transfusion - Staplers for Parynchyma	
	T_Struct	Yes/No	Transfusion - Staplers for Structures	
1	T_Glue	Yes/No	Transfusion - Fibrin Glue	
8	T_RFA	Yes/No	Transfusion - RFA	
	T_Cry	Yes/No	Transfusion - Cryoablation	
	ICUdays	Number	Days in ICU	
11	Infection	Yes/No	PO - Wound Infection	
8	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
VoResEvalDate	Date/Time	No Resection - Decision Date
VoResNoHandle	Yes/No	No Resection - Couldn't Handle Proposed Treatment
VoResRefused	Yes/No	No Resection - Refused Treatment
VoResMagnitude	Yes/No	No Resection - Magnitude Not Worth Benefits
VoResCeliacInvolve	Yes/No	No Resection - Celiac Trunk Involvement
VoResSMAInvolve	Yes/No	No Resection - SMA Involvement
VoResHepaticInvolve	Yes/No	No Resection - Hepatic Involvement
VoResIVCInvolve	Yes/No	No Resection - Inferior Vena Cava Involvement
VoResSMVInvolve	Yes/No	No Resection - SMV Involvement
VoResPVInvolve	Yes/No	No Resection - Portal Vein Involvement
VoResCirrhosis	Yes/No	No Resection - Cirrhosis
VoResMetastatic	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 Inf	ormation]	- 🗆 ×
Eile Edit View Insert Form	Nat Records Tools Window Help Adobe PDF Type a question for help Image:	• - 8 ×
Colorectal Cancer Follow-up Informati	on Add Record Find Record Delete Record	
MR Date of Visit ECOG performance status:	Follow-up Window:	
CEA: Albun CA19-9: Total	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 Experiment	
Status: © Died Death Date: [© N.E.D. © A.W.D., method of detection:	🗖 Lab 🗖 Radiologic evidence 🗖 Clinical evidence	
Record: H I Form View	▶★ of 1	

Figure 129: Colorectal Cancer Follow-Up Form

Field Name	Data Type	Description
8 ID	AutoNumber	
MR	Text	Meditech Medical Record Number for Patient
FUWin	Number	Follow-Up Windows
VisitDate	Date/Time	Visit Date
Weight	Number	Weight (bs.)
QOLscore	Number	QoL Score (0-100)
ECOG	Number	ECOG Score (0-4)
LabCEA	Number	Laboratory - CEA
LabCA19-9	Number	Laboratory - CA19-9
LabAlb	Number	Laboratory - Albumin
LabBili	Number	Laboratory - Bilirubin
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
SxChola	Yes/No	Symptoms - Cholangitis
SxBC	Yes/No	Symptoms - Biliary Colic
SxNau	Yes/No	Symptoms - Nausea
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
SxAbd	Yes/No	Symptoms - Abdominal Pain
SxBack	Yes/No	Symptoms - Back Pain
SxOT	Yes/No	Symptoms - Other
SxOTSpe	Text	Symptoms - Other - Specify
Status	Number	Status (NED, AWD, Died)
DeathDate	Date/Time	Death Date
StatusAWDLab	Yes/No	AWD - Lab Evidence
StatusAWDRad	Yes/No	AWD - Radiology Evidence
StatusAWDCli	Yes/No	AWD - Clinical Evidence
	10.004000	

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 Forma	at Records Tools Window Heln Ad	obe PDF	Type a guestion for help	
		▲ ++ *< ☎ ጬ / - ≡ <u>⊅</u> - <u>▲</u> - <u>⊿</u> - <u>-</u>	Ĩ. □·.	Mai (17 77
Breast Cancer S	creening			
atient Information				
Demographics:				
MR	Geography			
Height (in.)	Ethnicity			
Weight (lbs.)	Birthplace			
QOL	Grew Up			
ECOG	Lived Longest			
Pack Years Quit Y	rs Ago Drug Use 🔳			
Enviro. Exposure 🔳 Other	Factors Other - Specify			
Enviro. Exposure Dother Medical/Family Histo Breast Cancer	Factors Other - Specify	- Other Cancer		
Enviro. Exposure D Other Medical/Family Histo Breast Cancer Mother Breast Dx D	Factors Other - Specify	- Other Cancer Father Dx		
Enviro. Exposure D Other Medical/Family Histo Breast Cancer Mother Breast Dx M Mother Breast Dx - Age	Factors Other - Specify OTY Other 1 Breast Dx Other 1 Breast Dx	- Other Cancer Father Dx Mother Dx		
Enviro. Exposure Definition of the other other of the other of the other of the other other other of the other o	Factors Other - Specify OTY Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx Other 2 Breast Dx	Other Cancer Father Dx Mother Dx Other Relation 1		
Enviro. Exposure Other Medical/Family Histo Breast Cancer Mother Breast Dx M Mother Breast Dx Age Sister 1 Breast Dx M Sister 1 Breast Dx - Age	Factors Other - Specify OTY Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx Other 1 Breast Dx Other 2 Breast Dx Other 1 Breast Dx Other 1 Breast Dx	Other Cancer Father Dx Mother Dx Other Relation 1 Related Dx		
Enviro. Exposure Definition of the other Medical/Family History Breast Cancer Mother Breast Dx Markowski Mother Breast Dx Age Sister 1 Breast Dx Age Sister 1 Breast Dx Age Sister 2 Breast Dx Markowski Sister 3 Breast Dx Markowski	Factors Other - Specify OTY Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx Other 1 Breast Dx Other 2 Breast Dx	- Other Cancer Father Dx Mother Dx Other Relation 1 Related Dx Other Relation 2		
Enviro. Exposure Other Medical/Family Histo Breast Cancer Mother Breast Dx II Mother Breast Dx - Age Sister 1 Breast Dx - Age Sister 1 Breast Dx - Age Sister 2 Breast Dx II Sister 2 Breast Dx - Age	Factors Other - Specify OTY Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx Other 1 Breast Dx Other 1 Breast Dx	Other Cancer Father Dx Mother Dx Other Relation 1 Related Dx Other Relation 2 Related Dx		
Enviro. Exposure Other Medical/Family Histe Breast Cancer Mother Breast Dx M Mother Breast Dx - Age Sister 1 Breast Dx - Age Sister 1 Breast Dx - Age Sister 2 Breast Dx - Age Daughter 1 Breast Dx - Age	Factors Other - Specify OTY Other 1 Breast Dx Other 1 Breast Dx Other 2 Breast Dx Other 1 Breast Dx Other 1 Breast Dx	Other Cancer Father Dx Mother Dx Other Relation 1 Related Dx Other Relation 2 Related Dx		
Enviro. Exposure Other Medical/Family Histe Breast Cancer Mother Breast Dx M Mother Breast Dx Age Sister 1 Breast Dx Age Sister 2 Breast Dx Age Daughter 1 Breast M Cancer Age Daughter 1 Breast M Daughter 1 Breast M D Daughter 1 Breast M D D D D D D D D D D D D D	Factors Other - Specify OTY Other 1 Breast Dx Other 2 Breast Dx Other 1 Breast Dx Other 1 Breast Dx	Other Cancer Father Dx Mother Dx Other Relation 1 Related Dx Other Relation 2 Related Dx		

Figure 131: Breast Cancer Screening Form

	Field Name	Data Type	Description
80	D	AutoNumber	ID
	MR	Text	MR
	DemMR	Text	Demographics - Meditech Medical Record Number for Patient
	DemHeight	Number	Demographics - Height in Inches of Patient
	DemWeight	Number	Demographics - Weight in Pounds of Patient at Admission
	DemQOL	Number	Demographics - QOL Score (0-100)
	DemECOG	Number	Demographics - ECOG Score (0-4)
-	DemSHCigarette	Yesjivo Numbor	Demographics - Social History - Ligarettes (significant use)
	DemSHCigaretteOuit	VesiNo	Demographics - Social History - Clogerettes - Patk Tears
	DemSHCigaretteOuitYrs	Text	Demographics - Social History - Cioarettes - Out? - Years
	DemSHAlcohol	Yes/No	Demographics - Social History - Alcohol (significant use)
	DemSHDrugUse	Yes/No	Demographics - Social History - Drug Use
	DemSHExposure	Yes/No	Demographics - Social History - Environmental Exposure
1.3	DemSHOther	Yes/No	Demographics - Social History - Other
-	DemSHUthers	Text	Demographics - Social History - Other - Specify Demographics - Etherity
	DemGeographyBP	Text	Demographics - Georgaphy - Birtholace
13	DemGeographyGU	Text	Demographics - Geography - Grew Up
	DemGeographLL	Text	Demographics - Geography - Lived Longest
	FamilyBreastMother	Yes/No	Family History - Mother Breast Dx
	FamilyBreastMotherAge	Number	Family History - Mother Breast Dx Age
	FamilyBreastSister1	Number	Family history - Sister Librast Dx Early History - Sister Librast Dx
	FamilyBreastSister2	YesiNo	Family History - Sister 2 Bread Dx
	FamilyBreastSister2Age	Number	Family History - Sister2 Breast Dx Age
13	FamilyBreastDaughter1	Yes/No	Family History - Daughter1 Breast Dx
	FamilyBreastDaughter1Age	Number	Family History - Daughter1 Breast Dx Age
-	FamilyBreastDaughter2	Yes/No	Family History - Daughter2 Breast Dx
-	FamilyBreastOther1	Vec/No	Family History - Daugitterz breast bx Age
	FamilyBreastOther1Age	Number	Family History - Other1 Breast Dx Age
	FamilyBreastOther2	Yes/No	Family History - Other2 Breast Dx
	FamilyBreastOther2Age	Number	Family History - Other2 Breast Dx Age
	FamilyFatherDx	Number	Family History - Father Dx
	FamilyMotherDx	Number	Family History - Mother Dx
-	FamilyOther1 FamilyOther1Dm	Number	Family Fistory - Outer1 Family History - Other1 Dy
	FamilyOther2	Text	Family History - Other2
	FamilyOther2Dx	Number	Family History - Other2 Dx
	PriorBiopDate1	Date/Time	Prior Breast Biopsy - Date1
	PriorBiopPlace1	Text	Prior Breast Biopsy - Place1
	PriorBiopType1	Text	Prior Breast Biopsy - Type1
	PriorBiopHindings1	Text Date/Time	Prior Breast Biopsy - Findings1
	PriorBiopPlace2	Tevt	Prior Breach Biopcy - Date2
	PriorBiopType2	Text	Prior Breast Biopsy - Type2
	PriorBiopFindings2	Text	Prior Breast Biopsy - Findings2
	ReproMenses	Number	Repro History - Menses Onset
_	ReproLastMenstrual	Date/Time	Repro History - Last Menstrual Date
	Repromeno i ype	Number	Repro History - Menopause Type
	ReproChildren	Number	Repro History - Children
	ReproAbortions	Number	Repro History - Abortions
	ReproAgeFirstBirth	Number	Repro History - Age of First Birth
	ReproAgeMenopause	Number	Repro History - Age of Menopause
_	ReproBreastFeeding	Number	Repro History - Breast Feeding
	ReproHormoneOral	Number	Repro History - Hormone - Oral Contraceptives
	ReproHormoneHRT	Number	Repro History - Hormone - Hormone Renarcement Therany
	ReproHormoneHolistic	Number	Repro History - Hormone - Holistic/Homeopathic
	ReproGeneticCounseling	Memo	Repro History - Genetic Counseling Notes
1	SxSympPrimary	Number	Initial Symptoms - Primary Symptom
	SxSympSecondary1	Number	Initial Symptoms - Secondary Symptom
-	SxSympSecondary2	Number	Initial symptoms - secondary symptom Tabial Symptoms - secondary Symptom
	SxBreastMassDiscoverv	Text	Initial Symptoms – Breast Mass – Discovery
	SxMastalgia	Yes/No	Initial Symptoms - Mastalgia
	SxMastalgiaType	Number	Initial Symptoms - Mastalgia - Type
	SXINDischarge	res/No Text	Initial Symptoms - Nipple Discharge
	SxAxillaryMass	Yes(No	Initial Symptoms - Nipple Oscillarge - Type Initial Symptoms - Axillary Mass
	SxRetraction	Yes/No	Initial Symptoms - Nipple/Skin Retraction
	SxSystemicPain	Yes/No	Initial Symptoms - Systemic - Pain
	SxSystemicWeightLoss	Yes/No	Initial Symptoms - Systemic - Weight Loss
	5x5ystemicWeightLossLbs	Number	Initial Symptoms - Systemic - Weight Loss Lbs.
	SxSystemicDySpRea	Ves/No	anicia symptoms - systemic - byspried Initial Symptoms - Systemic - Pain
	SxSystemicOTSpe	Memo	Initial Symptoms - Systemic - Other
	WorkUpScreening	Yes/No	Workup - Screening
	WorkUpScreeningType	Number	Workup - Screening - Type
	WorkUpBreastMass	Yes/No	Workup - Breast Mass
	WorkUppreastMassSizeX	Number	Warkup - Breast Mass - Size - Width (cm) Warkup - Breast Mass - Size - Height (cm)
	WorkUpMammo	Number	Workup - Mammogram
	WorkUpBiRad	Number	Workup - BIRAD
	WorkUpUltra	Number	Workup - Ultrasound
-	WorkUpMRI RicoBalENIA	Number Vec/No	Workup - MRI Ricogu - Balcoble - Eine Maadle Accivation
-	BiopPalFNA BiopPalFNAIme	YesiNo	uiupsy - napatie - nite weede Aspiration Bionsy - Paloable - Fine Needle Aspiration - Imanion
	BiopPalCNB	Yes/No	Biopsy - Palpable - Core-cutting Needle Biopsy
	BiopPalCNBImg	Yes/No	Biopsy - Palpable - Core-cutting Needle Biopsy - Imaging
_	BiopPalIncB	Yes/No	Biopsy - Palpable - Incisional Biopsy
-	biopPalIncBImg BiopPalEvcB	res/No Vec/No	biopsy - Paipapie - Incisional Biopsy - Imaging Biopsy - Paipable - Excisional Biopsy
1	BiopPalExcBImg	Yes/No	Biopsy - Palpable - Excisional Biopsy - Imaging
	BiopNonPalMammo	Yes/No	Biopsy - Non-palpable - Mammography
	BiopNonPalUltra	Yes/No	Biopsy - Non-palpable - Ultrasound
	BiopDisp	Text	Biopsy - Disposition
	CBiopPalFNA	Yes/No	Contra Biopsy - Palpable - Fine Needle Aspiration
-	CBiopPalFNAImg	res/No Vec/No	Contra biopsy - Maipable - Hine Needle Aspiration - Imaging Contra Biopsy - Palpable - Core-cutting Needle Biopsy
-	CBiopPalCNBImm	Yes/No	Contra Biopsy - I alpable - Core-cutting Needle Biopsy - Imagina Contra Biopsy - Palpable - Core-cutting Needle Biopsy - Imagina
	CBiopPalIncB	Yes/No	Contra Biopsy - Palpable - Incisional Biopsy
	CBiopPalIncBImg	Yes/No	Contra Biopsy - Palpable - Incisional Biopsy - Imaging
	CBiopPalExcB	Yes/No	Contra Biopsy - Palpable - Excisional Biopsy
	CBiopMonPalExcB1mg	res/No Vec/No	Contra Biopsy - Paipable - Excisional Biopsy - Imaging Contra Biopsy - Non-nalnable - Mammorranby
	CBiopNonPalUltra	Yes/No	Contra Biopsy - Non-palpable - Ultrasound
	CBiopDisp	Text	Contra Biopsy - Disposition
	5 67 67		

Figure 132: Breast Cancer Screening Table Schema

File Edit View Insert Format Recon	ds Tools	Window Help Adobe	PDF	Type a guestion for help	
					100 J
	1 200 1 2				
Aridi y] -	-
Breast Cancer Stagi	ıg				
MR Date of Examination		- Alesander			
Physical Examination - Ipsalate	eral				
Breast Mass Size (cm) by		Skin Changes	lipple Changes		
Clock Dis. from Areola (cm)		Erythema	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	Туре			
Supraclav Nodes Mode Size (cm)	by	Type			
Physical Examination Control	atoral				
Breast Mass Size (cm) m	aterai				
Clock Dis from Areola (cm)		Skin Changes Endhema	Retraction		
		Edema 🔲	Discoloration		
Consistency					
Consistency		Dimpling	Erosion		
Consistency Skin Fixation		Dimpling 🔲 Sat. Nodes 🕅	Erosion Discharge		
Consistency Skin Fixation		Dimpling Sat. Nodes Ulceration	Erosion E		
Consistency Skin Fixation Pectoral Fixation Chest Wall Fixation Other Features Extra Nipples Arm Edema		Dimpling III Sat. Nodes III Ulceration III	Erosion U Discharge U		
Consistency Skin Fixation Pectoral Fixation Chest Wall Fixation Other Features Extra Nipples Axillary Nodes Node Size (cm)	by	Dimpling III Sat. Nodes III Ulceration III Type	Erosion U Discharge U		

Figure 133: Breast Cancer Staging Form

2	Field Name Data Type Description		
81	ID	AutoNumber	
	MR	Text	MR
1	Date	Date/Time	Date of Examination
333	Ipsa_Mass_SizeX	Number	Physical Exam - Ipsa - Breast Mass - SizeX
	Ipsa_Mass_SizeY	Number	Physical Exam - Ipsa - Breast Mass - SizeY
0	Ipsa Mass Clock	Number	Physical Exam - Ipsa - Breast Mass - Clock
2	Ipsa_Mass_AreolaD	Number	Physical Exam - Ipsa - Breast Mass - Areola Distance
32	Ipsa_Mass_Consistency	Text	Physical Exam - Ipsa - Breast Mass - Consistency
	Ipsa_Mass_FixSkin	Yes/No	Physical Exam - Ipsa - Breast Mass - Skin Fixation
0	Ipsa_Mass_FixPect	Yes/No	Physical Exam - Ipsa - Breast Mass - Pectoral Fixation
2	Ipsa_Mass_FixChest	Yes/No	Physical Exam - Ipsa - Breast Mass - Chest Wall Fixation
230	Ipsa_Skin_Erythema	Yes/No	Physical Exam - Ipsa - Skin Changes - Erythema
	Ipsa_Skin_Edema	Yes/No	Physical Exam - Ipsa - Skin Changes - Edema
0	Ipsa Skin Dimpling	Yes/No	Physical Exam - Ipsa - Skin Changes - Dimpling
2	Ipsa Skin Satellite	Yes/No	Physical Exam - Ipsa - Skin Changes - Satellite Nodes
8.	Ipsa Skin Ulceration	Yes/No	Physical Exam - Iosa - Skin Changes - Ulceration
	Ipsa Nipple Retraction	Yes/No	Physical Exam - Ipsa - Nipple Changes - Retraction
1	Ipsa Nipple Discolor	Yes/No	Physical Exam - Iosa - Nipple Changes - Discoloration
2	Ipsa Nipple Erosion	Yes/No	Physical Exam - Ipsa - Nipple Changes - Erosion
3.	Iosa Nipple Discharge	Yes/No	Physical Exam - Iosa - Nipple Changes - Discharge
	Insa Extra Ninnles	Yes/No	Physical Exam - Insa - Extra Ninnles
-	Insa Nodes Axil	Yes/No	Physical Exam - Inca - Axillary Nodes
35	Insa Nodes Axil SizeX	Number	Physical Exam - Insa - Axillary Nodes - SizeX
	Ipsa Nodes Axil SizeY	Number	Physical Exam - Ipsa - Axillary Nodes - SizeY
	Ipsa Nodes Axil Type	Number	Physical Exam - Iosa - Axillary Nodes - Type
	Ipsa Nodes Clav	Yes/No	Physical Exam - Ipsa - Supraclavicular Nodes
í.	Ipsa Nodes Clav SizeX	Number	Physical Exam - Iosa - Supraclavicular Nodes - SizeX
	Ipsa Nodes Clav SizeY	Number	Physical Exam - Ipsa - Supraclavicular Nodes - SizeY
	Insa Nodes Clay Type	Number	Physical Exam - Tosa - Supraclavirular Nodes - Type
	Iosa Arm Edema	Yes/No	Physical Exam - Inca - Arm Edema
8	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
8	Contra Mass Consistency	Text	Physical Exam - Contra - Breast Mass - Consistency
	Contra Mass EixSkin	YesíNo	Physical Exam - Contra - Breast Mass - Skin Eixation
	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
í.	Contra Skin Ervthema	Yes/No	Physical Exam - Contra - Skin Changes - Erythema
	Contra Skin Edema	YesíNo	Physical Exam - Contra - Skin Changes - Edema
	Contra Skin Dimpling	Yes/No	Physical Exam - Contra - Skin Changes - Dimoling
	Contra Skin Satellite	Yes/No	Physical Exam - Contra - Skin Changes - Satellite Nodes
Ŕ	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
Ŕ	Contra_Nipple_Discharge	Yes/No	Physical Exam - Contra - Nipple Changes - Discharge
	Contra_Extra_Nipples	Yes/No	Physical Exam - Contra - Extra Nipples
3	Contra_Nodes_Axil	Yes/No	Physical Exam - Contra - Axillary Nodes
	Contra_Nodes_Axil_SizeX	Number	Physical Exam - Contra - Axillary Nodes - SizeX
8	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
100	Contra_Nodes_Clav_SizeY	Number	Physical Exam - Contra - Supraclavicular Nodes - SizeY
	Contra_Nodes_Clav_Type	Number	Physical Exam - Contra - Supraclavicular Nodes - Type
	Contra_Arm_Edema	Yes/No	Physical Exam - Contra - Arm Edema
	Imaging_Study_1	Number	Imaging - Study - 1
10	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_SizeY	Number	Imaging - Study - 2 - SizeY
1	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
10	Imaging_Study_4	Number	Imaging - Study - 4
1	Imaging_Study_4_SizeX	Number	Imaging - Study - 4 - SizeX
	Imaging_Study_4_SizeY	Number	Imaging - Study - 4 - SizeY
	Imaging_Study_5	Number	Imaging - Study:- 5
100	Imaging_Study_5_SizeX	Number	Imaging - Study - 5 - SizeX
1	Imaging_Study_5_SizeY	Number	Imaging - Study - 5 - SizeY
	Primary_Stage	Number	Staging - Primary Tumor Stage
	Node_Stage	Number	Staging - Lymph Node Stage
80	Mets_Stage	Number	Staging - Metastatic Stage
-	Prelim_Outlook	Number	Preliminary Outlook
			L

Figure 134: Breast Cancer Staging Table Schema

	Date of Admission				
Operativ Breast Sid	e <u>Details</u> e	Resection Type	Success	×	
Surgical	Reconstruction				
Prosthet	c Type Tissue/Pr	osthetic Type Tissue A	ione Type		
Bleeding	Infection Seror	na 🔳			
Other De	e <mark>tails</mark> se ₪ Duration (hrs	s.)			
Drain U Port-a-(Cath Placement 🔳	Additional Organ Biopsie	S		

Figure 135: Breast Cancer Resection Form

Field Name	Data Type	Description
	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 💌	-
- ■ ♥ ● ▷ ♥ ※ ☜ ☜ ♡ % 2+ \$+ ♥ ┓♥ ₩ ▸ ♥ ₪	1 to 10 + 12 +	
Arial • 9 • B <i>I</i> <u>U</u> ≣ ≣ ≣ <u>2</u> • <u>A</u> •		
Breast Cancer Chemotherapy		
MR Date of First Treatment Date of Last Treatment		
Treatment Course		
Chemo Type		
AC CMF DosACI TAC		
FEC E FAC Other Specify		
Include Herpacin?		
<u>Growth Factors</u>		
Port-a-Cath Placed?		
Complete Therapy?		
Complications		
Dose Dense? Cognitive Changes Hospitalization Febrile Neutro Emesis Cardiac		
Neuronathy Amenorches Other Spacify		
rd: 11 + 1 + 11 + 1 of 1 (Filtered)		

Figure 137: Breast Cancer Chemotherapy Form

Field Name	Data Type	Description
₽ND	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?
	1 () () () () () () () () () (

Figure 138: Breast Cancer Chemotherapy Table Schema

Erosoft Access - [Breast File Edit View Insert	Radiation Format Records Tools	Window Help Adobe PDF	Type a question for help	- 5
• 🖬 🖏 🎒 🗟 💖	み 略 色 や 色 2 - 9 - 3	↓↓↓♥┓♥₩ ₩ ₩ 8 <i>I</i> ⊻ ≡ ≡ ≡ <u>♪</u> ↓ <u>↓</u> .	⊡ ′a • Q . ∠ • □• □• .	
Breast Canc	er Radiother: st Treatment Date of Last	apy		
Sites Treated				
Breast Treatment 🗐 Whole Breast Dose	Total # Fractions	Dose/Fraction		
Tumor Bed Boost Dose	Total # Fractions	Dose/Fraction		
<u>Chest Treatment</u>	Total # Fractions	Dose/Fraction		
Scar Boost 🔳 Dose	Total # Fractions	Dose/Fraction		
Axilla Treatment III Dose	Total # Fractions	Dose/Fraction		
Supraclavicular Fossa I Dose	reatment III Total # Fractions	Dose/Fraction		
rd: 14 4 1	▶ ▶ ▶ ★ of 1 (Filtered)			
			FLTR	

Figure 139: Breast Cancer Radiotherapy Form

2511	Field Name	Data Type	Description
81	ID	AutoNumber	
1111	MR	Text	MR
ŝ	Date First	Date/Time	Date of First Treatment
	Date Last	Date/Time	Date of Last Treatment
	Breast Treatment	Yes/No	Breast - Treatment
1	Breast_Dose	Number	Breast - Whole Breast Dose
ŝ	Breast_Fractions	Number	Breast - Total # Fractions
	Breast_FracDose	Number	Breast - Dose/Fraction
	Breast_Boost	Yes/No	Breast - Tumor Bed Boost
î.	Breast_Boost_Dose	Number	Breast - Tumor Bed Boost - Dose
8	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
1	Chest Dose	Number	Chest - Whole Breast Dose
ŝ	Chest_Fractions	Number	Chest - Total # Fractions
	Chest_FracDose	Number	Chest - Dose/Fraction
	Chest_Boost	Yes/No	Chest - Scar Boost
1	Chest Boost Dose	Number	Chest - Scar Boost - Dose
ŝ	Chest Boost Fractions	Number	Chest - Scar Boost - Total # Fractions
	Chest_Boost_FracDose	Number	Chest - Scar Boost - Dose/Fraction
	Axilla_Treatment	Yes/No	Axilla - Treatment
1	Axilla_Dose	Number	Axilla - Whole Breast Dose
8	Axilla_Fractions	Number	Axilla - Total # Fractions
	Axilla_FracDose	Number	Axilla - Dose/Fraction
	Supra_Treatment	Yes/No	Supraclavicular Fossa - Treatment
1	Supra_Dose	Number	Supraclavicular Fossa - Whole Breast Dose
ŝ	Supra_Fractions	Number	Supraclavicular Fossa - Total # Fractions
	Supra_FracDose	Number	Supraclavicular Fossa - Dose/Fraction
	IntNodes_Treatment	Yes/No	Internal Mammary Nodes - Treatment
1	IntNodes_Dose	Number	Internal Mammary Nodes - Whole Breast Dose
8	IntNodes_Fractions	Number	Internal Mammary Nodes - Total # Fractions
	IntNodes_FracDose	Number	Internal Mammary Nodes - Dose/Fraction
	Unsched_Interrupt	Yes/No	Unscheduled Treatment Interruption
1	Unsched_Interrupt_Days	Number	Unscheduled Treatment Interruption - Days
8	Unsched_Interrupt_Why	Text	Unscheduled Treatment Interruption - Why?
	Complete	Yes/No	Patient Completion
0	Complete_IfNoWhy	Number	Patient Completion - If No Why
1	Concur_Chemo_TAM	Yes/No	Concurrent TAM or Chemo?
8	Toxic_RT	Yes/No	Toxicity of RT
	Toxic_RT_Grade	Number	Toxicity of RT - Grade

Figure 140: Breast Cancer Radiotherapy Table Schema

Microsoft Access - [Breast_Mets]		_ 🗆 ×
E Elle Edit View Insert Format Records Iools Window Help Adobe PDF	Type a question for help	• _ & ×
₩ • ₩ 월 월 Q ♥ % ħ 6 ♡ % 월 ¥ ¥ Ў Ѣ ▼ ₩ ₩ % 6 ⊡ m • Q .		
▼ ▼ B / U 臣喜君 <u>久</u> ・ <u>人</u> ・ <u>/</u> ・□・	Ţ	
Breast Cancer Overt Metastatic Treatment		
Dieuse Cancer Overe inclustatic incluiment		
MR		
Palliative Measures		
Dalliative Surgers 2 C Dalliative Partiation? C		
"Decimenta Tara		
Palliative Chemotherapy?		
Other Pall, Measures?		
Endocrine Therapy		
Antiestrogens Progestins # # Regimens		
Aromatase Inhibitors 🔳 Bisphosphanates 🗐		
Other Specify		
n ne benefation de la contraction de la Contraction de la contraction de la cont		
Record: 1 > > > of 1 (Filtered)		
PLIK PLIK	<u>n av at ar a</u>	/

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

File Edit View	preast_FU]					_0
The Try Tour	Insert Form	at <u>R</u> ecords <u>T</u> e	ools <u>W</u> indow <u>H</u> elp Ado <u>b</u>	e PDF	Type a question for help	
2 • 日 🖻 🖨 [a 💖 🐰 I	h R 0 0	1 21 XI 🦻 🏹 6		2.	
- Aria	l)	• 9	• B <i>I</i> <u>U</u> = = =	<u>∆</u> • <u>∠</u> • <u></u> •		
Broget Ce	meor I	T wollo				in to so
Di east Ca	incel 1	onow-c	ЪЪ			
MR Date	e of Visit	FU Window				
			•			
Performance	Status					
ECOG Score QoL	Score					
Serum Marke	rs					
CEA	CA15-3					
A 10-11-11-11-11-11-11-11-11-11-11-11-11-1	Datasta	576 Mar 100	Allester			
Albumin	Bilitabili	and the state of the state of	Aikaline			
and the second se	ndings	Tumor Sta	itus			
Mammographic Fi						
Mammographic Fi						
Mammographic Fi	oped Sym	ptoms	- I have been be created			
Mammographic Fi New/Redevel New Primary? Ru	oped Sym ecurrent Dise	<u>ptoms</u> ase in Ipsalater	al New Disease in Contral	ateral		
Mammographic Fi New/Redevel New Primary? Ru New Cancer in Ot	oped Sym ecurrent Dise her High-Risk	<u>ptoms</u> ase in Ipsalater Organ Metasta	al New Disease in Contral	ateral		
Mammographic Fi New/Redevel New Primary? Ra New Cancer in Ot	oped Sym ecurrent Dise her High-Risk	ptoms ase in Ipsalater Organ Metasta	al New Disease in Contral	ateral Leukemia		
Mammographic Fi	oped Sym ecurrent Dise her High-Risk	ptoms ase in Ipsalater Organ Metasta	al New Disease in Contral	ateral Leukemia		
Mammographic Fi	oped Sym ecurrent Dise her High-Risk	ptoms ase in Ipsalater Organ Metasta	al New Disease in Contral	ateral Leukemia		
Mammographic Fi New/Redevel New Primary? R New Cancer in Ot Other Specify Patient Status	oped Sym ecurrent Dise her High-Risk	ptoms ase in Ipsalater Organ Metasta	al New Disease in Contral	ateral Leukemia		
Mammographic Fi New/Redevel New Primary? R New Cancer in Ot Other Specify Patient Status	oped Sym ecurrent Dise her High-Risk	ptoms ase in Ipsalater Organ Metasta	al New Disease in Contral	ateral Leukemia		
Mammographic Fi New/Redevel New Primary? R New Cancer in Ot Other Specify Patient Status	oped Sym ecurrent Dise her High-Risk	ptoms ase in Ipsalater Organ Metasta	al New Disease in Contral	aterai Leukemia		
Mammographic Fi New/Redevel New Primary? Ru New Cancer in Ot Other Specify Patient Status Status	oped Sym ecurrent Dise her High-Risk	ptoms ase in lpsalater organ Metasta m m Radiological I	al New Disease in Contral	aterai Leukemia		
Mammographic Fi New/Redevel New Primary? Re New Cancer in Ot Other Specify Patient Status Status Clinical Detection ord: 14	ecurrent Dise her High-Risk	ptoms ase in lpsalater organ Metasta	al New Disease in Contral	aterai Leukemia		

Figure 143: Breast Cancer Follow-Up Form

Field Name	Data Type	Description
P 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

		76.7	
MR Patholog	y Date		
Specimen Type		Diagnosis	
	• 1.000/2012/01/10/10		
Precancerous Sp	ecimen Charact	eristics	
Precancerous Type			
La della d	<u>•</u>		
Ductal Carcinoma	a Insitu (DCIS) S	pecimen Characteristics	
Histology	Nuclear Grade		
Comedo Necrosis	Size	Multifocality	
	-		
# of DCIS Slides	Margin	Calcification	
		Specimen Characteristics	
Lobular Carcinon		opeemen enalactensues	
Lobular Carcinon	Size	Necrosis	
Lobular Carcinon	Size	Necrosis	
Lobular Carcinon Histology Margin	Size		

Figure 145: Breast Cancer Pathology Form

6	Field Name	Data Type	Description
8	D	AutoNumber	
117	MR	Text	MR
1	Date	Date/Time	Pathology Date
£.	Specimen_Type	Number	Specimen Type
	Diag_Type	Number	Diagnosis Type
11	Precan_Type	Number	Precancerous Type
1	DCIS_Hist	Number	DCIS Histology
í.	DCIS_Grade	Number	DCIS Grade
1	DCIS_Comedo	Number	DCIS Comedo Necrosis
11	DCIS_Size	Number	DCIS Size
4	DCIS_Multifoc	Number	DCIS Multifocality
	DCIS_Slides	Number	DCIS Slides
	DCIS_Margin	Number	DCIS Margin
11	DCIS_Calc	Number	DCIS Calcification
1	LCIS_Hist	Number	LCIS Histology
	LCIS_Size	Number	LCIS Size
	LCIS_Necrosis	Number	LCIS Necrosis
11	LCIS_Margin	Number	LCIS Margin
4	LCIS_Calc	Number	LCIS Calcification
6	DIC_Type	Number	DIC Type
1	LIC_Type	Number	LIC Type
0	LIC_Grade	Number	LIC Grade
1	LIC_Size	Number	LIC Size
8	LIC_LVI	Number	LIC Lymphovascular Invasion
1	LIC_Necrosis	Number	LIC Necrosis
10	LIC_Margin	Number	LIC Margin
11	LIC_Skin	Number	LIC Skin Involvement
<u>(</u>	LIC_Nipple	Number	LIC Nipple Involvement
	LIC_IS_Presense	Number	LIC IS Presense
1	LIC_IS_Type	Number	LIC IS Type
1	LIC_IS_Grade	Number	LIC IS Grade
	LIC_IS_Invasion	Number	LIC IS Invasion
2	LIC_IS_EIC	Number	LIC IS EIC
11	LIC_Node_Type	Number	LIC Node Type
11	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
8	LIC_Rec_Prog	Number	LIC Progesterone Receptor
14	LIC_Rec_IHC	Number	LIC IHC Receptor
U.	LIC_Rec_FISH	Number	LIC FISH Receptor
1	LIC_Stage	Number	LIC Pathology Stage
	02 0000 00		5 20/X 30

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
Т3	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

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)
А	Class Discretization: Discrete target classes are
	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
	are discretized via MDL method. Derived from Data Set A.
С	Correlation-based Feature Selection: Attribute subsets are chosen
	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
Nouvel Network	MLP	Multi-layer Perceptron
Ineural Inetwork	RBF	Radial Basis Function
Regression	LNG	Linear Regression

Table 21: Regression Algorithms

Clinical Factors - 1	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)
F	Unaltered Data Set: Uses original
Ľ	instance data with numeric target classes.
Г	Correlation-based Feature Selection: Attribute subsets are chosen
Г	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 - 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	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	Size T	-Test	- ML	Algo	rithms	s vs. L	ogist	tic Re	gress	ion	251		
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

fumor Size T-Test - ML Algs. w/ AdaBoostM1 vs. Logistic Regression													
R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1				
	-	-	-	E	=	=	=	=	=				
	=	=:	=	=	=	=	=	=	=				
	=	=	=	=	E	=	=	=	=				
	=	=	=	Ξ	=	=	=	=	=				
	(AB1	LAB1 J48.AB1 - = = = =	AB1 J48.AB1 IB1.AB1 = = = = = = =	AB1 J48.AB1 IB1.AB1 IB2.AB1 = = = = = = = = = =	AB1 J48.AB1 IB1.AB1 IB2.AB1 IB3.AB1 - - - = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	AB1 J48.AB1 IB1.AB1 IB2.AB1 IB3.AB1 LWL.AB1 - - - = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	AB1 J48.AB1 IB1.AB1 IB2.AB1 IB3.AB1 LWL.AB1 NVB.AB1 - - - = </td <td>X.AB1 J48.AB1 IB1.AB1 IB2.AB1 IB3.AB1 LWL.AB1 NVB.AB1 BN1.AB1 - - - =</td> <td>X.AB1 J48.AB1 IB1.AB1 IB2.AB1 IB3.AB1 LWL.AB1 NVB.AB1 BN1.AB1 BN2.AB1 - - - =</td>	X.AB1 J48.AB1 IB1.AB1 IB2.AB1 IB3.AB1 LWL.AB1 NVB.AB1 BN1.AB1 - - - =	X.AB1 J48.AB1 IB1.AB1 IB2.AB1 IB3.AB1 LWL.AB1 NVB.AB1 BN1.AB1 BN2.AB1 - - - =				

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

Tumor Size 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	=	-		-	Ξ	=	E	н	=	=		
В	=	=	=	H	=	=	=	=	=	=		
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)								
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 - J	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	=	=	=	=	=	=	=	= 5	=	=	=
D	Ŧ	=	=	=	=	Ŧ	=	=	=	=	=

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

T-Stag	e T-Tes	st - ML	Algs.	w/ Ada	aBoost	M1 vs.	Logistic	Regre	ssion				
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	=	=	=	=	=	=	=	=	=	=			
+ : Super	· : Superior to LGR = : Equivalent to LGR - : Inferior to LGR												

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

T-Stage	e T-Te	st - MI	_ Algs	. w/ B	aggin	g vs. L	ogistic	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	+	+	+	+	+	+	+	+	+	+
С	=	=	=	=2	=	=	=	=	=	=
D	=	Ξ.	H	-	=	=	3H)	=	2H	=
+ : Superi	or to LG	R =:Ec	uivalen	t to LGR	- : Infe	rior to LC	R	- 1 C	177 71	jun Ac

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-9	Stage	l)								
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	icatior	1 - N-SI	tage - A	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	÷	Ξ	=	+	+	÷	+	+	=	+	=
В	Ŧ	=	=	=	=	#2	=	# 3	=	i , ≥	=
С	=	=	=	=	=	=8	=	=	=	=	=
D	=	=	=	=	=	Ŧ	=	=	=	=	=

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

N-Stag	e T-Te	st - ML	Algs.	w/ Ad	laBoos	tM1 vs	. Logist	ic Regr	ression			
Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1		
A	=	Ξ.	+	+	=	+	=	=	=	=		
В	Ŧ	=	Ξ	=	ŧ.	=	=	÷.	=	=2		
С	=	=	=	=	=3	=	=	=	=	=		
D	=	=	Ξ.	=	=	=	=	=	=	=		
+ : Super	: Superior to LGR = : Equivalent to LGR - : Inferior to LGR											

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

N-Stage	+ T-Tes	t - ML	Algs.	. w/ Ba	igginç	jvs. L	.ogistic	Regre	ession			
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	=	=	=	=	=	=	=	=	=	=		
D	=	=	=	=	=	=	=	=	=	=		
+ : Superio	: Superior to LGR = : Equivalent to LGR - : Inferior to LGR											

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
A	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	40. SI				
Data Set	LGR	1R.AB1	J48.AB1	IB1.AB1	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
А	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

Vascula	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	=	=	=	=	=	=	=	=	=	=	=				
+ : Superio	+ : Superior to LGR = : Equivalent to LGR - : Inferior to LGR														

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

Vascul	ature T-	Test -	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	=	=	=	Ξ	200	=	=	Ξ	=	=
С	=	=	=	=	Ξ.	=	= :	=	=	=
D	÷	=	=	=	Ξ	=	=	=	=	=
+ : Superi	or to LGR	= : Equiv	alent to l	GR -:	Inferior t	o LGR	8	10. 	92 2	2

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

Vascul	ature T-	Test - I	/asculature T-Test - ML Algs. w/ Bagging vs. Logistic Regression													
Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BO	NVB.BG	BN1.BG	BN2.BG	BN3.BG						
A	=	(=)	=	# 3	=	=	=	=	=	Ξ						
В	=	=	=	=	=	=	1) 1)	=	÷.							
С	1=	=	=	= 8	=	=	= 8	=	=	=						
D	=	=	=	=	=	=	=	=	Ŧ	2						

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	Classification - Histology														
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			

Figure 183: Histology - Accuracy Results (Perc
--

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	Classification - Histology - 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	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	-	=	=	(=)	=	T S	=	=	() 	# 0	=
В	T:	12 H	=	=	=	#	=	=	=	=	=
С	49	=	=	=	=	= 8	=	+	+	=	=
D		=	=	=	=	=	=	+	+	÷.	=

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

Histold	ogy T-T	'est - N	IL Algs	s. w/ A	daBoo	stM1 v	s. Logi	stic Reg	gressio	n
Data Set	1R.AB1	J48.AB1	IB1.AB1	IB2.AB	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	=	±۵	=	=	н	=	=	=	=	=
В	Ξ.	=	÷.	=	=	=	H	=	=	=
С	=	=	=	=	=	=	=	=	=	=
D	=	=	=:	=	=	=	=	=	=	—
+ : Super	ior to LG	₹ = : Eq	uivalent t	LGR	- : Inferio	r to LGR				.

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

listology T-Test - ML Algs. w/ Bagging vs. Logistic Regression														
1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BC	NVB.BG	BN1.BG	BN2.BG	BN3.BG					
=	# 8	= 5	=	=	= 3	=	=	=	÷.					
=	=	н	=	=	Ξ	=	=	=	=					
-	=	=	=	=	+ :	+	±_	=	=					
=	=	=	=	=	+	+	÷.	=	=					
1	y 1-1 IR.BG = = =	y I-lest - N IR.BG J48.BG = = = = = = = = = = = =	y I-Test - WL Aig IR.BG J48.BG IB1.BG = = = = = = = = = = = = = = = =	y I-lest - IVIL Algs. W/ IR.BG J48.BG IB1.BG IB2.BG = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	y I-lest - ML Aigs. W/ Baggi IR.BG J48.BG IB1.BG IB2.BG IB3.BG = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	y I-lest - WL Algs. W/ Bagging vs IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BC = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = + = = = = = = + = +	y 1-lest - ML Algs. W/ Bagging vs. Logis IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BQ NVB.BG =	y 1-lest - ML Algs. W/ Bagging vs. Logistic Reg IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BQ NVB.BG BN1.BG = <td>y 1-lest - ML Algs. W/ Bagging vs. Logistic Regression IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BQ NVB.BG BN1.BG BN2.BG =</td>	y 1-lest - ML Algs. W/ Bagging vs. Logistic Regression IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BQ NVB.BG BN1.BG BN2.BG =					

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	+ : Superior to LGR = : Equivalent to LGR - : Inferior to LGR											

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

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	=	=	=	Ξ.		=	=	=	=	=	
С	=	=	=	=	=	=	=	=	=	=	
D	=	=	=	=	2	=	=	=	=	=	
+ : Super	F : Superior to LGR = : Equivalent to LGR - : Inferior to LGR										

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	=	+	=	=	=	=	=	=	=	=	
В	=	Ξ	=	=	E.	=		=	=	E.	
С	=	=	=	=	=	=	=	=	=	=	
D	=	=	=	=	=	=	÷	=	=	=	
+ : Superi	+ : Superior to LGR = : Equivalent to LGR - : Inferior to LGR										

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.

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

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

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	=	=	=	=	=	=	=	=	=	=	=		
D	=	=	=	=	=	=	=	=	=	=	=		
+ : Superior to LGR = : Equivalent to LGR - : Inferior to LGR													

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

ECOG 6-Month 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	=	E.	E)	=	=	=	=	E	=	=		
В	=	÷.	H	E	=	=	=	E	E	=		
С	=	=	=	=	= :	=	=:	=	=	=		
D	=	=	=	=	±	=	= <	=	(=	(±)		
F: Superior to LGR = : Equivalent to LGR - : Inferior to LGR												

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

ECOG 6	ECOG 6-Month T-Test - ML Algs. w/ Bagging vs. Logistic Regression													
Data Set	1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	BN1.BG	BN2.BG	BN3.BG	NVB.BG				
А	=	=	=	±.	=	=	Ξ	Ξ	=	=				
В	=	=	=	=	=	=	=	=	=	=				
С	=:	=	=	(=)	=	=	= :	# ()	=	=				
D	=	=	=	=	=	=	2	=	=	=				

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	icatio	n - 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	=	=	=	=	=	+	+	=	=	=
В	+) #	=	=	=	=	=	=	=	=	=
С	E.	=	=	=	=	=	=	=	=	=	=
D	=	=	=	=	=	=	=	=	=	=	=

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

ECOG	9-Mon	th T-Te	st - M	L Algs	. w/ Ad	aBoost	M1 vs.	Logistic	Regre	ssion		
Data Set	1R.AB1	J48.AB1	IB1.AB	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1		
A	=	=	=	=	=	+	+		=	۲		
В	=	=	=	=	=	=	=	E	=	Ξ.		
С	=	=	=	= :	=	Ξ.	=	=	=	=		
D	=	=	=	2 0	=	=	=	=	=	=		
+ : Super	: Superior to LGR = : Equivalent to LGR - : Inferior to LGR											

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

ECOG 9-Month T-Test - ML Algs. w/ Bagging vs. Logistic Regression												
1R.BG	J48.BG	IB1.BG	IB2.BG	IB3.BG	LWL.BG	NVB.BG	BN1.BG	BN2.BG	BN3.BG			
-	=	=	=	=	+	+	=	= 3	=8			
	216	=		=	=		=	=	=			
-	=	=	=	=:)	=	=	=	=	=			
=	÷ i	=	=	=	=	÷	=	=	=			
1	IR.BG	IR.BG J48.BG = = = = = = = =	IR.BG J48.BG IB1.BG = = = = = = = = = = = = = = = = = = = = =	INDITITI 1-Test - ML Aigs IR.BG J48.BG IB1.BG IB2.BG = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	INDITITI 1-16St - ML Aigs. W/ B IR.BG J48.BG IB1.BG IB2.BG IB3.BG = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	INDITITI 1-16St - ML Algs. W/ Dagging IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BG = = = = + = = = = + = = = = = = = = = = = = = = = = = = = = = = = = =	INDITITI 1-16St - ML Algs. W/ Dagging VS. L(IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BG NVB.BG = = = = + + = = = = + + = = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	INDITITI 1-16St - ML Aigs. W/ Bagging VS. Logistic IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BG NVB.BG BN1.BG = = = = + + = = = = = + + = = = = = + + = = = = = = = = = = = = = = = = = = = = = = = = = = = = =	INDITITI 1-16St - ML AIGS. W/ Dagging VS. Logistic Regre IR.BG J48.BG IB1.BG IB2.BG IB3.BG LWL.BG NVB.BG BN1.BG BN2.BG = = = = + + = = = = = = + + = = = = = = + + = = = = = = = = = = = = = = = = = = = = = = = = = = = = =			

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	Classification - ECOG 12-Month - AdaBoostM1														
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

ECOG	12-M	onth	T-Tes	t - ML	. Algo	rithm	s vs. I	ogis	tic Re	gress	sion
Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	=	=	=	E.	=		=	=	=	(=)	=
B	=	=	=	=	=	=	=	=	=	=	=
C	=	=	=	=	=	=	=	=	=	=	=
D	=	=	=	=	=	=	=	=	=	=	=

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

ECOG	12-Mo	nth T-T	est -	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	=	=	=	=	=	=	=	=	=	z
В	3 11 2	2H%	=	Ξ.	E.	=	=	5H6	E	
С	=	=	=	=	=	=	=	=	=	=
D	Ξ	=	=	=	=	=	=	2	=	=
+ : Super	or to LGF	R = : Equ	ivalent	to LGR	: Inferior	to LGR		5	8	6

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	=	=	=:	=	=	=	=	=	=	=
В	=	=	H	=	1	23	=	E	Ξ	
С	=	=	= 8	=	=	=	=	=	=	=
D	=	=	=	=	=	=	=	Ξ.	Ξ.	Ξ.

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	ı - 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

Surviv	al T-1	Fest -	ML A	lgorit	hms ۱	/s. Lo	gistic	Regr	essio	n	
Data Set	ZR	1R	J48	IB1	IB2	IB3	LWL	NVB	BN1	BN2	BN3
A	E.	2	=	=	—	(a)	=) = s	=)=:) = 2
В	=	E.	=	=	=	=	=	=	Ē	=	=
С	=	=	=	=	=	=	=	+	=	= :	=
D	=	=	=	=	=	()=	=	+	=	=	=

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

Surviv	al T-Te	st - ML	Algs	. w/ Ad	aBoos	tM1 vs.	Logisti	c Regre	ssion	
Data Set	1R.AB1	J48.AB1	IB1.AB	IB2.AB1	IB3.AB1	LWL.AB1	NVB.AB1	BN1.AB1	BN2.AB1	BN3.AB1
A	=	=	=:	Ξ	=	=	Ξ	=	=	=
В	=	=	=	=	=	=	2	=	=	H
С	=	=	= 8	=	=	=	=	=	=	=
D	=	=	=	=	=	=	94 14 25	=	=	=
+ : Superi	ior to LGF	R = : Equ	ivalent	to LGR	: Inferior	to LGR	2	15	5.	5

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

Survival 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	=	=	2	=	=	=	=	=	=	=		
В		=	Æ	=	=	=	=	=	=	=		
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 - 1	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 - T	umor Si	ze - Ba	gging ar	nd Stack	king			
Data Set	LNR	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK	1
E	0.27	0.29	0.36	0.28	0.03	0.26	0.24		0.15
F	0.41	0.45	0.30	0.42	0.17	0.41	0.38	5	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

Tumor Algorit	Size hms	T-Test - vs. Line	ML ar Reg	ression	14	
Data Set	ZR	M5M	M5R	RBF	MLP2	MLP3
E	-	=	=	-0	.	(- -3
F	-	=	1	<u>i</u> 201	i an	24
⊢ +:Super	ior to Ll	NR =:Equ	ivalent to	LNR -:	- Inferior to	LNR

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

Tumor w/ Met	Size T- a-Learn	Test - N ers vs.	IL Algo Linear I	rithms Regress	ion		
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK
E	=	=	=	5	=	=	8
F	=	-	æ	-	=	E.	-
+ : Super	ior to LNR	= : Equiv	alent to LM	NR -:Infer	ior to LNR	- E	

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

ECOG Algorit	6-Mo hms	nth T-Te vs. Line	est - ML ar Reg	- ression	Î.	
Data Set	ZR	M5M	M5R	RBF	MLP2	MLP3
E		+	=	=	+	+
F	-	1 H2 1	=	-	1	=
+ : Super	ior to Ll	NR = : Equ	uivalent to	LNR -:	Inferior to	LNR

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

ECOG	ECOG 6-Month T-Test - ML Algorithms										
w/ Meta-Learners vs. Linear Regression											
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.

Regres	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	0.00				
F	0.04	0.13	0.14	0.08	0.25	0.16	0.13	0.06				

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 Algorit	ECOG 9-Month T-Test - ML Algorithms vs. Linear Regression										
Data Set	ZR	M5M	M5R	RBF	MLP2	MLP3					
E	=	+	+	Ш	=	=					
F - = - = = =											
+ : Superio	+ : Superior to LNR = : Equivalent to LNR - : Inferior to LNR										

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

ECOG 9	ECOG 9-Month T-Test - ML Algorithms										
w/ Meta	w/ Meta-Learners vs. Linear Regression										
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.

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	Regression - ECOG 12-Month - Bagging and 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

ECOG Algorit	12-Mo hms	onth T-1 vs. Line	ſest - N ar Reg	IL ression	1	
Data Set	ZR	M5M	M5R	RBF	MLP2	MLP3
E	1 51	17	175)	-	+	+
F	20	1	-0	-	±≤	=
+ : Super	ior to Ll	NR =: Equ	uivalent to	LNR -:	Inferior to	LNR

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

ECOG 12-Month T-Test - ML Algorithms w/ Meta-Learners vs. Linear Regression								
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK	
E	.	=	=	÷	+	+	E.	
F	=	<u>Le</u>	+	=	=	=	2	
+ : Super	ior to LNR	= : Equiv	alent to LN	R : 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.

Regression - Survival									
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

Regression - Survival - Bagging and Stacking										
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		0.04	

Figure 262: Survival - R-Squared Results - AdaBoostM1



Figure 263: Survival - Regression Results Graph



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

Surviv Algorit	al T-T thms '	ˈest - ML vs. Line	- ar Reg	ression		
Data Set	ZR	M5M	M5R	RBF	MLP2	MLP3
E		=	(T 5)	=	+	+
F	-	1	-0	-	±=≤	=
+ : Super	ior to Ll	NR = : Equ	uivalent to	LNR :	Inferior to	LNR

Figure 265: Survival - T-Test vs. Linear Regression
Survival T-Test - ML Algorithms w/ Meta-Learners vs. Linear Regression											
Data Set	M5M.BG	M5R.BG	LNR.BG	RBF.BG	MLP2.BG	MLP3.BG	STK				
E	=	-	+	÷	+	+	E.				
F	44	<u>1</u>	=	40	=	=	2				
+ : Superior to LNR = : Equivalent to LNR - : Inferior to LNR											

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)

								3.5	ਰ ਗ	Con	fus:	ion	Ma	tr	ix ===
a	b	с	d	е	£	g	h	i	Ĵ.	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	c = Ductal_Adeno_of_Pancreas
1	0	1	2	0	0	0	1	0	0	0	0	0	0	I.	d = Neuroendocrine_(Islet_Cell/Carcinoid)
0	0	0	1	0	0	0	0	0	0	0	0	0	0	T.	e = Von_Hippel-Lindau_Syndrome
0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	f = Duodenal_Adenocarcinoma
0	0	0	0	0	1	0	0	0	0	0	0	0	0	Ĩ.	g = Distal_Cholangiocarcinoma
0	0	0	1	0	0	0	0	1	0	0	1	0	0	T	h = Renal_Mets
0	0	1	0	0	0	0	0	1	0	0	2	0	0	Ť	i = Cystadenoma
0	0	0 C	0	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	I.	n = Benign_Cyst

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

Pro	bability Distribution Table For EU	SNoNode				×					
	Histology	Presum	ptiveDx	FALSE	TRUE						
Aden	ocarcinoma_of_Pancreas/NOS	Pancreat	ic_Tumor	0.92	9 0.0	71					
Aden	ocarcinoma_of_Pancreas/NOS	Periampul	lary_Tumor	0.62	5 0.3	75					
Aden	ocarcinoma_of_Pancreas/NOS	Suspicious_P	ancreatic_Cyst	0.	5 (0.5					
Aden	ocarcinoma_of_Pancreas/NOS	IPMT	(IPMN	0.	5 (0.5					
Aden	ocarcinoma_of_Pancreas/NOS	Suspicious_Bile	e_Duct_Stricture	0.7	5 0.	25					
Aden	ocarcinoma_of_Pancreas/NOS	Otl	her	0.	5 (0.5					
A	mpullary_Adenocarcinoma	Descent of	te Arcasea Note de la composite dal	0.7							-
A	mpullary_Adenocarcinoma	Probability	Distribution Tabl	e For EU	SStaging		_				<u>~</u>
A	mpullary_Adenocarcinoma		Histology	i	EUSPortal	T4	T3	T2	T1	ΤX	
A	mpullary_Adenocarcinoma	Adenocarcino	oma_of_Pancrea	s/NOS	TRUE	0.368	0.368	0.158	0.053	0.053	*
A	mpullary_Adenocarcinoma	Adenocarcino	oma_of_Pancrea	s/NOS	FALSE	0.128	0.333	0.385	0.128	0.026	
A	mpullary_Adenocarcinoma	Ampullar	y_Adenocarcinon	าล	TRUE	0.429	0.143	0.143	0,143	0.143	
E	Ouctal_Adeno_of_Pancreas	Ampullar	y_Adenocarcinon		FALSE	0.048	0.619	0.143	0,143	0.048	
1)uctal_Adeno_of_Pancreas	Ductal_A	deno_of_Pancrea	as	TRUE	0.2	0.2	0.2	0.2	0.2	
E	Ouctal_Adeno_of_Pancreas	Ductal_A	deno_of_Pancrea	as 📗	FALSE	0.085	0.119	0.729	0.017	0.051	
	A Probability Distribution Table	For CTNodeOm	i -			XI	0.2	0.2	0.2	0.2	
			TOUT	1			0.067	0.6	0.2	0.067	
	Histology	SXPru _	TRUE	1	ALSE		0.2	0.2	0.2	0.2	
leur	Adenocarcinoma_of_Pancreas/	NOS FALSE	0.3	1	4	J.7 🔺	0.143	0.429	0,143	0.143	1
leur	Adenocarcinoma_of_Pancreas	NOS TRUE	0.045		0.9	55	0.2	0.2	0.2	0.2	
leur	Ampullary_Adenocarcinoma	FALSE	0.562	1	0.4	38	0.333	0.333	0,111	0.111	
leur	Ampullary_Adenocarcinoma	a TRUE	0.167		0.8	33	0,2	0.2	0.2	0.2	
	Ductal_Adeno_of_Pancreas	s FALSE	0.241		0.7	59	0.143	0.429	0,143	0.143	
	Ductal_Adeno_of_Pancreas	S TRUE	0.25		0.	75	0.2	0.2	0.2	0.2	
	Neuroendocrine_(Islet_Cell/Carc	inoid) FALSE	0.1			0.9	0.091	0.455	0.273	0.091	
	Neuroendocrine_(Islet_Cell/Carc	inoid) TRUE	0.25		0.	75	0.2	0.2	0.2	0.2	
	Von_Hippel-Lindau_Syndron	ne FALSE	0.25		0.	75	0.077	0.692	0.077	0.077	
	Von_Hippel-Lindau_Syndron	ne TRUE	0.5			0.5	0.2	0.2	0.2	0.2	:
	Duodenal_Adenocarcinoma	a FALSE	0.167		0.8	33	0.143	0.429	0,143	0.143	i
	Duodenal_Adenocarcinoma	a TRUE	0.5	1		1.5	0.2	0.2	0.2	0.2	
	Distal_Cholangiocarcinoma	FALSE	0.5			1.5	0.143	0.429	0,143	0.143	-
	Distal_Cholangiocarcinoma	a TRUE	0.25		0.	75					
	Renal_Mets	FALSE	0.125		0.8	/5					
	Renal_Mets	TRUE	0.5			1.5					
	Cystadenoma	FALSE	0.7			J.3					
	Cystadenoma	TRUE	0.5			7.5					
	MEN-I	FALSE	0.25		0.	75					

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 ===

а	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	c = Ductal_Adeno_of_Pancreas
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	k = Pseudopapillary_Tumor
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 SxSatiety(2): 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-		
	${\tt Eweka.attributeSelection.CfsSubsetEval-Sweka.}$		
	attributeSelection.BestFirst -D 1 -N 5		
Instances:	72		
Attributes:	s: 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		
Thresh	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.