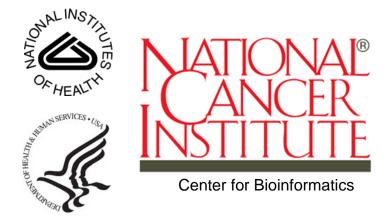
I-SPY USER'S GUIDE

Version 1.0



CREDITS AND RESOURCES

I-SPY Development and Management Teams		
Development	Documentation	Product and Program Management
Subha Madhavan	Laura Jackel	Subha Madhavan
Alex Jiang	Ying Long	
Kevin Rosso	Huaitian Liu	
Ryan Landy		
Himanso Sahni		
David Bauer		
Huaitian Liu		
Michael Harris		
Ram Bhattaru		
Ye Wu		
Ying Long		
Don Swan		
Dana Zhang		
Vesselina Bakalov		
Nick Xiao		
Gregg Silk		
¹ Science Application International Corporation (SAIC)	² National Cancer Institute Center for Bioinformatics (NCICB)	³ National Cancer Institute

I-SPY Development and Management Teams		
Development	Documentation	Product and Program Management
⁴ University of Maryland, Department of Diagnostic Radiology	⁵ VA Maryland Healthcare System	⁶ MIRC Committee, Radiological Society of North America (RSNA)
⁴ Northern Taiga Ventures, Inc.		

Contacts and Support		
NCICB Application Support	http://ncicb.nci.nih.gov/NCICB/support Telephone: 301-451-4384 Toll free: 888-478-4423	

TABLE OF CONTENTS

About This Guide	V
Purpose	v
Audience	
Topics Covered	
Text Conventions Used	vi
Chapter 1	
About I-SPY 1.0	1
About I-SPY	1
About I-SPY Functions	
Chapter 2	
Getting Started with I-SPY 1.0	3
Launching I-SPY	
Creating a User Account	
Logging In	
Accepting I-SPY Provisions	
Welcome to I-SPY 1.0	
Getting Help	
Application Support	
Logging Out	, 0
Chapter 3	
Conducting an Identifier Lookup	9
ID Lookup Overview	9
Looking Up a Patient or Sample Identifier	
Displaying/Hiding the Patient Sample Information	11
Downloading Patient Sample Information to an Excel File	12
Creating a PatientDID List with the ID Lookup	12
Chapter 4	
Conducting Searches	13

Search Overview	13
Performing a Clinical Query	13
Performing an IHC Level of Expression Query	16
Performing an IHC Loss of Expression Query	18
Chapter 5	
High Order Analysis	21
High Order Analysis Overview	21
Performing a Class Comparison	
Performing a Principal Component Analysis	
Performing Hierarchical Clustering Analysis	
Performing Correlation Scatter Plot Analysis	
Performing Categorical Plot Analysis	30
Chapter 6	
Viewing Results	33
Results Overview	33
Search Results	33
Clinical Reports	34
IHC Level of Expression Search Results	36
IHC Loss of Expression Search Results	36
High Order Analysis Results	37
Class Comparison Report	37
Principal Component Analysis Plot	40
Hierarchical Clustering Report	
Correlation Scatter Plot	43
Categorical Plot Analysis	
Chapter 7	
Managing Lists	45
Managing Lists Overview	45
PatientDID Lists	
Gene Lists	
Viewing the Data Items in a List	
Removing Data Items to Create a New List	
Combining Existing Lists to Create a New List	
Deleting an Entire List	
Adding a New "Custom" List	
Glossary	51

ABOUT THIS GUIDE

This section introduces you to the *I-SPY User's Guide*. It includes the following topics:

- Purpose on page v
- Audience on page v
- Topics Covered on page v
- Text Conventions Used on page vi

Purpose

This guide provides an overview of I-SPY. This book is organized into chapters that parallel I-SPY's workflow.

Audience

This guide is designed to assist researchers and investigators using the I-SPY Analysis Portal application.

Topics Covered

If you are new to I-SPY, read this brief overview, which explains what you will find in each chapter.

- Chapter 1 provides an overview of the I-SPY program.
- Chapter 2 provides instructions to start using I-SPY.
- Chapter 3 describes how to search on patient or sample identiers. The results show the patients that fulfill the criteria.
- Chapter 4 describes how to perform a clinical and an IHC (Immunohistochemistry) Level of Expression and Loss of Expression query.
- Chapter 5 extends the basic knowledge of the previous chapters and shows you
 how to work with five types of high order analyses.
- Chapter 6 describes how to view all the results generated from searches and high order analyses.
- Chapter 7 describes how to manage user-defined and study-defined patient and gene identifier lists.

Text Conventions Used

The following table explains conventions used in this guide. The various typefaces represent interface components, keyboard shortcuts, toolbar buttons, dialog box options, and text that you type.

Convention	Description	Example
Bold	Highlights names of option buttons, check boxes, drop-down menus, menu commands, command buttons, or icons.	Click Search .
URL	Indicates a Web address.	http://domain.com
text in SMALL CAPS	Indicates a keyboard shortcut.	Press ENTER.
text in SMALL CAPS + text in SMALL CAPS	Indicates keys that are pressed simultaneously.	Press SHIFT + CTRL.
Italics	Highlights references to other documents, sections, figures, and tables.	See Figure 4.5.
Italic boldface monospace type	Represents text that you type.	In the New Subset text box, enter Proprietary Proteins.
Note:	Highlights information of particular importance	Note: This concept is used throughout the document.
{ }	Surrounds replaceable items.	Replace {last name, first name} with the Principal Investigator's name.

Table Documentation conventions

CHAPTER ABOUT I-SPY 1.0

This chapter introduces you to I-SPY and provides an overview of I-SPY functions. Topics in this chapter include:

- About I-SPY on page 1
 - About I-SPY Functions on page 2

About I-SPY

The NCI Center for Bioinformatics (NCICB), in collaboration with physicians, researchers, and cooperative groups, has designed I-SPY. Clinical trials are critical to identifying markers and mechanisms of resistance in therapy, and I-SPY is a multicenter clinical trial for women undergoing neoadjuvant chemotherapy from breast cancer. I-SPY is a web-based system which supports correlative data analysis and centralized reporting of results to catalyze the transition from uniform to tailored care.

I-SPY facilitates collaboration, provides an infrastructure for data management, analysis and communication, and develops a commitment to sharing information and developing data standards.

About I-SPY Functions

Users can perform a variety of tasks in I-SPY. Table 1.1 describes each I-SPY task.

Task	Description	
Perform a Patient or Sample Identifier Lookup	Search the database for patient or sample identifiers. Display, download, and save the data associated with the search criteria. For more information, see <i>Conducting an Identifier Lookup</i> on page 9)	
Perform a Search	Perform one of the following types of queries: Clinical query IHC Level of Expression query IHC Loss of Expression query For more information, see <i>Conducting Searches</i> on page 13.	
Perform a High Order Analysis	Run the following types of higher order analyses: Class comparisons Hierarchical clustering Principal component analyses Correlation scatter plot Categorical plot analysis. For more information, see <i>High Order Analysis</i> on page 21.	
View Results	View Search and High Order Analysis results. For more information, see <i>Viewing Results</i> on page 33.	
Manage Lists	Manage user-defined or study-defined patient or gene identifier lists. You can use them to filter queries or perform analysis. For more information, see <i>Managing Lists</i> on page 45).	

Table 1.1 I-SPY user tasks

CHAPTER

2

GETTING STARTED WITH I-SPY 1.0

This chapter introduces general I-SPY procedures and how to obtain help to use I-SPY. Topics in this chapter include:

- Launching I-SPY on page 3
- Creating a User Account on page 4
- Logging In on page 5
- Accepting I-SPY Provisions on page 5
- Welcome to I-SPY 1.0 on page 6
- Getting Help on page 7
- Application Support on page 7
- Logging Out on page 8

Launching I-SPY

To launch I-SPY, follow these steps:

1. Go to the I-SPY portal on the NCICB website:

http://ispy-analysis-stage.nci.nih.gov



The I-SPY login page appears (Figure 2.1).

Figure 2.1 I-SPY login page

powered by

Creating a User Account

Each I-SPY user is given a unique username and password. The user name and password you are assigned determines your access rights for the software. To set up a user account, you must:

- Contact NCICB Application Support:
 - NCICB@pop.nci.nih.gov
 - 888-478-4423 (toll-free) or 301-451-4384 (local)

OR

 Go to the NCICB I-SPY login page and click the request username/password link to send an e-mail requesting a username and password to NCICB Application Support.

Logging In

To log into I-SPY, you need the username and password assigned to you by the I-SPY Administrator.

1. On the login page, enter your **username** and **password**.



Figure 2.2 I-SPY login

2. Click the **Submit** button. If your login is successful, the Legal Rules of the Road page appears (*Figure 2.3*).

Accepting I-SPY Provisions

Once you log in, the Legal Rules of the Road page appears. After reading the provisions, click the **CLICKING HERE** link (*Figure 2.3*) in the lower right-hand corner.

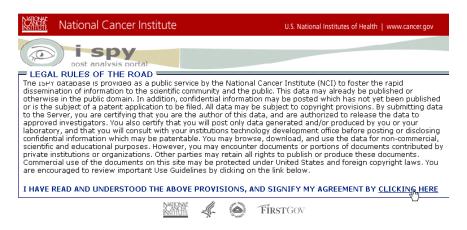


Figure 2.3 Legal Rules of the Road page

The I-SPY workspace appears (Figure 2.4).

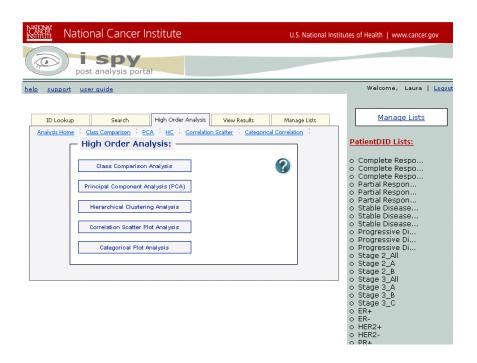


Figure 2.4 The I-SPY workspace

Welcome to I-SPY 1.0

The I-SPY workspace comprises a set of five tabs, a blue panel, help links, and a logout link. The five tabs enable you to perform the following functions:

- 1. Perform an ID lookup.
- 2. Perform complex searches.
- 3. Perform higher order analyses.
- 4. View results of the searches and analyses.
- 5. Manage lists.

The blue panel displays pre-defined patient identifier lists and the default gene identifier list. As you add your own lists, the new lists will appear in red.

Getting Help

Information about how to use I-SPY is easily accessed from I-SPY's menu (*Figure 2.5*) in the top left of the I-SPY workspace.



Figure 2.5 I-SPY's menu

Table 2.1 describes each item on the I-SPY toolbar.

Help	How to Access
Complete online help	To access the complete version of online I-SPY help, click the help link located in the I-SPY menu.
	For complete page-level help, click ? on any I-SPY page. For brief field help, click ? .
Application support	To obtain support for I-SPY, click the support link located under the I-SPY menu.
User's Guide	To access a pdf version of the <i>I-SPY User's Guide</i> , click the user guide link located under the <i>I-SPY</i> menu.

Table 2.1 Getting help with I-SPY

Application Support

You can find additional support at the NCICB Applications Support Web site. To access the site, do the following:

Click the **support** link in the I-SPY menu. The NCICB Applications Support Group page appears.

Logging Out

To log out of I-SPY, follow these steps.

1. On the I-SPY workspace, click the **logout** link (*Figure 2.6*) in the upper right-hand corner.



Figure 2.6 Logout link

You are logged out of I-SPY.

CHAPTER 3

CONDUCTING AN IDENTIFIER LOOKUP

This chapter describes how to use I-SPY to look up patient identifiers or sample identifiers.

Topics in this chapter include:

- ID Lookup Overview on page 9
- Looking Up a Patient or Sample Identifier on page 10

ID Lookup Overview

The ID Lookup function enables you to find information about samples for a given patient by entering either sample or patient identifiers. Once you perform the lookup, you can also perform the following tasks:

- Displaying/Hiding the Patient Sample Information
- Downloading Patient Sample Information to an Excel File
- Creating a PatientDID List with the ID Lookup

Note: You can use a PatientDID list to filter I-SPY queries and perform data analysis.

Looking Up a Patient or Sample Identifier

When you search for a patient identifier, I-SPY displays the patient along with all the samples associated with the patient. If you search for a sample identifier, I-SPY displays the patient associated with the sample identifier. To perform an ID lookup, follow these steps:

1. From the ID Lookup page (*Figure 3.1*), enter a valid patient identifier, such as 1001, or enter a valid sample identifier, such as 209512.

Note: To enter multiple identifiers, separate the identifiers with commas. For example, enter 1001, 1002.

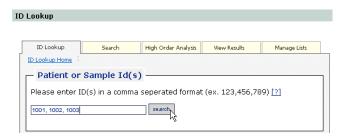


Figure 3.1 Entering identifiers

2. Click the Search button.

The patients associated with the identifier(s) appear below the **Search** button (*Figure 3.2*).

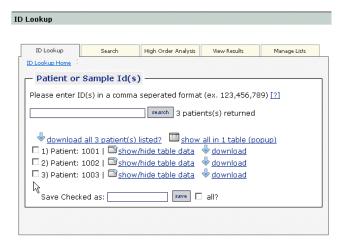


Figure 3.2 Found patients

Displaying/Hiding the Patient Sample Information

Once you perform an ID Lookup, you can display sample information for either an individual patient or multiple patients. Follow these steps:

1. To display all the samples collected for an individual patient click in ext to the patient's row.

The table highlights the lookup criteria (Figure 3.3).



Figure 3.3 Patient table data

Table 3.1 describes the sample information associated with the patient.

Item	Special Instructions	
ISPY ID	The identifier for the patient.	
LabTrak ID	The identifier for the sample collected.	
Timepoint	 T1 T2: Early treatment day 1, cycle 2 T3: Inter-regimen T4: Prior to surgery for response evaluation forma and sample post-surgery 	
Core Type	The type of substance used to store the sample.	
Section Info		

Table 3.1 Understanding the patient table data page

Note: To hide the table, data click .

2. To display data for multiple patients in one table click above the list of patients.

All the patients' data are shown in one table listed in descending order by patient identifier.

Note: To hide the table data, close the window.

Downloading Patient Sample Information to an Excel File

From the ID Lookup page, you can download one patient's sample data to a file or download all the listed patients' data to the same file. Follow these steps.

To download an individual patient's data to an Excel file, follow these steps:

- 1. Click next to the patient for which you want to download data.
- 2. Name the file and select a location.

The individual patient's data is saved to the Excel file.

If you searched for *multiple* patients, to save all the patients' data to the same Excel file, follow these steps.

- 1. Click above the list of patients to download all the patient sample information to one file.
- Name the file and select a location.
 All the patients' data is saved to the same Excel file listed in descending order by patient identifier.

Creating a PatientDID List with the ID Lookup

From the ID Lookup page, you can save multiple patients' data to a PatientDID list (see *PatientDID Lists*). You can use PatientDID lists to further filter a query or analyze data. To create an I-SPY PatientDID list, follow these steps.

1. Select the box next to each patient to be saved to the PatientDID list (*Figure 3.4*) or select the **All** box to select all of the patients.

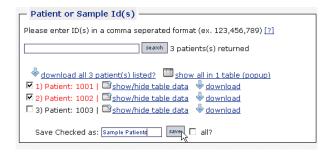


Figure 3.4 Saving to a PatientDID list

- 2. Name the list.
- 3. Click the save button.

The list name appears in red in the blue panel at the bottom of the PatientDID Lists.

Note: To further modify the new PatientDID list, see *Managing Lists*.

CHAPTER 4

CONDUCTING SEARCHES

This chapter describes how to perform queries to generate results, such as reports and plots.

Topics in this chapter include:

- Search Overview on page 13
- Performing a Clinical Query on page 13
- Performing an IHC Level of Expression Query on page 16
- Performing an IHC Loss of Expression Query on page 18

Search Overview

The Search function enables you to perform queries from the following categories:

- Performing a Clinical Query
- Performing an IHC Level of Expression Query
- Performing an IHC Loss of Expression Query

Report results are listed on the View Results page.

Performing a Clinical Query

A *clinical query* enables you to generate clinical reports using customized search criteria. The search criteria filter the report based on clinical, MR, or pathology parameters. For example, you can create a clinical query that finds patients between the ages of 31 and 50 and had a complete response within timepoints T1 and T2.

To define a clinical query, follow these steps:

1. On the Clinical Query Form page, you are required to fill in at least one search criteria (*Figure 4.1*).

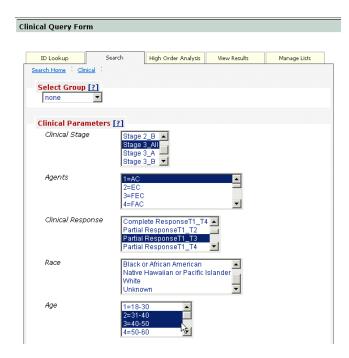


Figure 4.1 Clinical Query Form page (top portion)

2. Table 4.1 lists the available search criteria:

Note: To select more than one option in a list box, SHIFT-click or CTRL-click.

Criteria	Item Name	Special Instructions
Select Group	Select Group	A <i>group</i> is a pre-defined or user-defined PatientDID list comprising patient identifiers with certain characteristics. Select a group to filter the query to a collection of patients. Lists that you created appear in red.
Clinical Parameters	Clinical Stage	Select a clinical stage to further filter the query: Stage 2_AII Stage 2_A Stage 2_B Stage 3_AII Stage 3_A Stage 3_B
	Agents	Select an agent to further filter the query.

Table 4.1 Clinical Query search criteria instructions

Criteria	Item Name	Special Instructions
	Clinical Response	Select one or more clinical responses and the appropriate timepoint range. Complete Response Partial Reponse Stable Disease Progressive Disease
	Race	Select one or more races.
	Age	Select one or more age ranges.
MR Parameters	Morphology	Select an MRI parameter to further filter the query based on the radiologist measurement.
	Percent LD	This group of options enables you to specify the percentage of LD (Longest Diameter) change in the size of the tumor between two timepoints.
		Select the timepoint range in which to analyze the percentage of LD change:
		 PERCENT_LD_CHANGE_T1_T2: The specified percentage change in LD occurred between timepoints T1 and T2.
		 PERCENT_LD_CHANGE_T1_T3: The specified percentage change in LD occurred between timepoints T1 and T3.
		 PERCENT_LD_CHANGE_T1_T4: The specified percentage change in LD occurred between timepoints T1 and T4.
		Select the greater than/equal to (>=) or less than/equal to (<=) indicator and enter the percentage of LD change to search for in the selected timepoint range.
Pathology	Pathology Tumor Size	Specify the tumor size and associated biomarkers to filter the query. Select the greater than/equal to (>=) or less than/equal to (<=) indicator and enter a value in centimeters of the tumor size.
	Status	Specify the Pathology Status:
		ER+: Estrogen receptor positive ER-: Estrogen receptor negative PR+: Progesterone receptor positive PR-: Progesterone receptor positive HER2+: HER2 positive HER2-: HER2 negative

Table 4.1 Clinical Query search criteria instructions

3. Once you fill in at least one search criteria, you are required to enter a name for the clinical query. The name must be unique among all the queries in the current session.

To clear all the entries on the page, click the Clear button.

4. To submit the guery and generate the Clinical report, click the **Submit** button.

The Clinical report will be listed on the View Results page. From the report, you can save all or some of the patients' data to an I-SPY PatientDID list. You can use PatientDID lists to further filter a query or perform analysis. See *Clinical Reports*.

Performing an IHC Level of Expression Query

An *IHC* (*Immunohistochemistry*) *Level of Expression* query enables you to filter a search with one or more timepoints, biomarkers, and stain characteristics. The report results list records that satisfy the specified search criteria.

To perform a IHC Level of Expression query, follow these steps:

1. On the IHC Level of Expression Query Form page, the following search criteria are available to filter the query (*Figure 4.2*).

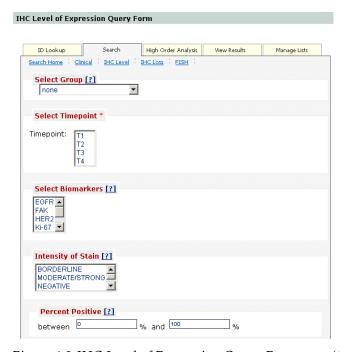


Figure 4.2 IHC Level of Expression Query Form page (top portion)

Table 4.2 lists the available search criteria:

Note: To select more than one option in a list box, SHIFT-click or CTRL-click.

Criteria	Special Instructions	
Select Group	A <i>group</i> is a pre-defined or user-defined PatientDID list comprising patient identifiers with certain characteristics. Select a group to filter the query to a collection of patients. Lists that you created appear in red.	
Select Timepoint	You are required to specify at least one timepoint during which the selected criteria are fulfilled. T1, T2, T3, T4	
Select Biomarkers	Select one or more biomarkers to filter the query:	
	 P27 Ki-67 EGFR CCND1 P53 HER2 BCL2 FAK 	
Intensity of Stain	Select an option that best describes the intensity of stain:	
	 Negative Borderline Weak Moderate_Strong Unevaluable 	
Percent Positive	Enter the percent positive range to filter the query.	
Localization of Stain	Select an option that best describes the localization of stain: None Membrane Nucleus Cytoplasm Membrane_and_Cytoplasm Nuclear_and_Cytoplasm NA or Not Applicable	
Distribution of Stain	Select an option that best describes the distribution of	
	NoneHomogenousHeterogenous	

Table 4.2 IHC Level of Expression Query search criteria instructions

- 2. Once you fill in the search criteria, you are required to enter a name for the IHC query. The name must be unique among all the queries in the current session.
 - To clear all the entries on the page, click the **Clear** button.
- 3. To submit the query and generate the IHC Level of Expression report, click the **Submit** button.

The IHC Level of Expression report will be listed on the View Results page. From the report, you can save all or some of the patients' data to an I-SPY PatientDID list. You can use a PatientDID list to further filter a query or perform analysis. For more information, see IHC Level of Expression Search Results.

Performing an IHC Loss of Expression Query

An *IHC* (*Immunohistochemistry*) Loss of Expression query enables you to filter a search with one or more timepoints, the P27 biomarker only, and invasive and benign range characteristics. The report results list records that satisfy the specified search criteria.

To perform a IHC Loss of Expression query, follow these steps:

1. On the IHC Loss of Expression Query Form page, the following search criteria are available to filter the query (*Figure 4.2*).

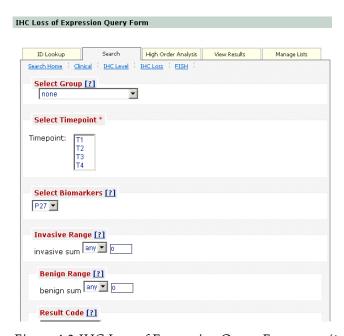


Figure 4.3 IHC Loss of Expression Query Form page (top portion)

Table 4.2 lists the available search criteria:

Note: To select more than one option in a list box, SHIFT-click or CTRL-click.

Criteria	Special Instructions
Select Group	A <i>group</i> is a pre-defined or user-defined PatientDID list comprising patient identifiers with certain characteristics. Select a group to filter the query to a collection of patients. Lists that you created appear in red.
Select Timepoint	You are required to specify at least one timepoint during which the selected criteria are fulfilled. T1, T2, T3, T4
Select Biomarkers	IHC Loss of Expression data is available only for biomarker P27 .
Invasive Range	Specify equal to, greater than, or less than to define the range invasive sum: • = • >= • <= Specify the value for the invasive sum.
Benign Range	Specify equal to, greater than, or less than to define the range benign sum: • = • >= • <= Specify the value for the benign sum.

Table 4.2 IHC Loss of Expression Query search criteria instructions

- Once you fill in the search criteria, you are required to enter a name for the IHC query. The name must be unique among all the queries in the current session.
 To clear all the entries on the page, click the Clear button.
- 3. To submit the query and generate the IHC Loss of Expression report, click the **Submit** button.

The IHC Loss of Expression report will be listed on the View Results page. From the report, you can save all or some of the patients' data to an I-SPY PatientDID list. You can use a PatientDID list to further filter a query or perform analysis. For more information, see *IHC Loss of Expression Search Results*.

CHAPTER

5

HIGH ORDER ANALYSIS

This chapter describes how to use I-SPY to run higher order analyses.

Topics in this chapter include:

- High Order Analysis Overview on page 21
- Performing a Class Comparison on page 22
- Performing a Principal Component Analysis on page 24
- Performing Hierarchical Clustering Analysis on page 26
- Performing Correlation Scatter Plot Analysis on page 27
- Performing Categorical Plot Analysis on page 30

High Order Analysis Overview

The High Order Analysis function enables you to perform the following analyses:

- Performing a Class Comparison
- Performing a Principal Component Analysis
- Performing Hierarchical Clustering Analysis
- Performing Correlation Scatter Plot Analysis
- Performing Categorical Plot Analysis

Report results are listed on the View Results page.

Performing a Class Comparison

A Class Comparison analysis alows you to identify genes and reporters that are differentially expressed between two groups. To perform a Class Comparisons, follow these steps:

1. The Class Comparison Analysis Form page (*Figure 5.1*) enables you to define the criteria to perform a class comparison.

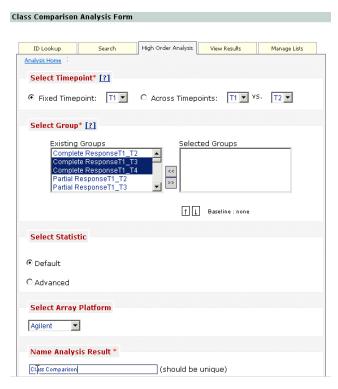


Figure 5.1 Class Comparison Analysis Form page

2. You are required to complete at least one criteria for the class comparison. *Table 5.1* lists the available criteria:

Criteria	Item Name	Special Instructions
Select Timepoint	Fixed Timepoint	Select a timepoint in which to perform the analysis. This option compares <i>two groups</i> at the same timepoint.
	Across Timepoints	Select a range of timepoints in which to perform the analysis. This option analyzes <i>one group</i> at different timepoints.

Table 5.1 Class Comparison criteria instructions

Criteria	Item Name	Special Instructions
Select Group	Existing Groups Selected Groups	A <i>group</i> is a pre-defined or user-defined PatientDID list comprising patient identifiers with certain characteristics. Select a group to filter the query to a collection of patients. Lists that you created appear in red. For a <i>Fixed Timepoint</i> analysis, select two groups (compares two groups at the same timepoint). For an <i>Across Timepoints</i> analysis, select one group
	Baseline	(analyzes one group at different timepoints).
	Baseline	For a Fixed Timepoint analysis, the baseline is determined by the second group in the Selected Groups box.
		For an Across Timepoints analysis, the baseline is determined by the first timepoint in the chosen range.
		The (baseline) appears in red next to your selection.
Select Statistic	Default	Select to perform a default statistical analysis.
	Advanced	Select to define additional statistical analysis options.
	Statistical Method	Select the appropriate statistical method:
		T-test: Two Sample Test to identify genes showing statistically significant differences between two samples.
		Wilcoxon Test: Man-Whitney Test is the non- parametric test analog to the independent two- sample t-test. This test is used in place of a two- sample t-test when the populations being compared are not normal.
	Multiple Comparison Adjustment	Family-wise Error Rate (FWER): Bonferroni False Discover Rate (FDR): Benjamini-Hochberg
	Fold Change	The default is >=2. Specify the threshold for the differential regulation. This returns differential expression ratios between tumor and non-tumor samples for a particular reporter.
	p-value	The probability for obtaining the differences in expression values between tumor (or a subtype of tumor) and non-tumor samples. The default is <=0.05.
Select Array Platform	Select Array Platform	Select the array platform.

Table 5.1 Class Comparison criteria instructions

3. You must enter a title/name for this analysis in the **Name Analysis Result** text box. This name must be unique among all your queries in this session.

4. To submit your criteria and create a Class Comparison report, click the **Submit** button.

Performing a Principal Component Analysis

A *Principal Component analysis* is a dimensionality reduction algorithm, which identifies clusters of samples that may have similar gene expression profiles. To perform a Principal Component Analysis, follow these steps:

1. The Principal Component Analysis (PCA) Form page (Figure 5.2) enables you to define criteria to perform a PCA.

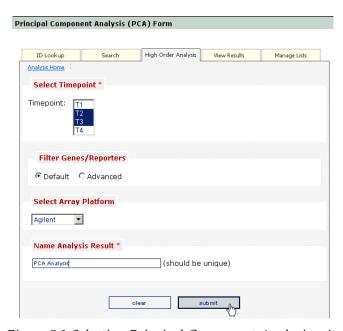


Figure 5.2 Selecting Principal Component Analysis criteria

2. You are required to complete at least one criteria for the Principal Component analysis. *Table 5.2* lists the available criteria:

Criteria	Item Name	Special Instructions
Select Timepoint	Timepoint	Select one or more timepoints in which to perform the analysis.
Filter Genes/Reporters	Default	Select to perform a default statistical analysis.
	Advanced	Select to define additional gene/reporter filters.
	Constrain reporters by variance (Gene Vector) percentile: %	Enter a percentage which selects the reporters whose variances of the log ratio (or Log2 signals) across all experiments were among the top percentile of variance of all reports identified. For example, 70% chooses reporters with the top 30 (100 - 70) percentile of variance.
	Constrain by GeneList	A Gene list is a pre-defined or user-defined list comprising genes with certain characteristics. Select a gene list to filter the query. Lists that you created appear in red. The default gene list is defaultGene1 .
Select Array Platform	Select Array Platform	Select an array platform.

Table 5.2 Principal Comparison Analysis criteria instructions

- 3. You must enter a title/name for this analysis in the **Name Analysis Result** text box. This name must be unique among all your queries in this session.
- 4. To submit your criteria and create a Principal Comparison Analysis report, click the **Submit** button.

Performing Hierarchical Clustering Analysis

Hierarchical Clustering analysis creates a dendrogram of the samples in the analysis. To perform a Hierarchical Clustering, follow these steps:

1. The Hierarchical Clustering Analysis Form (*Figure 5.3*) enables you to fill in criteria for a hierarchical clustering.

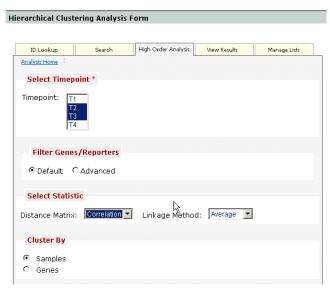


Figure 5.3 Selecting Hierarchical Clustering criteria

2. You are required to enter at least one step for the hierarchical clustering. *Table* 5.3 lists the available criteria:

Criteria	Item Name	Special Instructions
Filter Genes/Reporters	Default	Select to perform a default statistical analysis.
	Advanced	Click to define additional gene/reporter filters.
	Constrain reporters by variance (Gene Vector) percentile: %	Enter a percentage which selects the reporters whose variances of the log ratio (or Log2 signals) across all experiments were among the top percentile of variance of all reports identified. For example, 70% chooses reporters with the top 30 (100 - 70) percentile of variance.
	Constrain by GeneList	A Gene list is a pre-defined or user-defined list comprising pgenes with certain characteristics. Select a gene list to filter the query. Lists that you created appear in red. The default gene list is defaultGene1 .

Table 5.3 Hierarchical Clustering criteria instructions

Criteria	Item Name	Special Instructions
Select Statistic	Distance Matrix	Select a distance matrix option: Correlation measures the relative shape of the gene regulations rather than the absolute levels. This is a natural choice, because it is widely used to measure gene correlations. Euclidean distance is the most common distance measure. It measures the absolute level of gene regulation.
	Linkage Method	Select a linkage option to affect the shape of the resulting clusters: • Average linkage is the average of all pair-wise distances between members of the two clusters. • Single linkage is the minimum distance between two clusters. • Complete linkage is the maximum distance between two clusters.
Cluster By	Cluster by	Leave the default to cluster on Samples or cluster by Genes .
Select Array	Select Array Platform	Select an array platform.

Table 5.3 Hierarchical Clustering criteria instructions

- 3. You must enter a title/name for this analysis in the **Name Analysis Result** text box. This name must be unique among all your queries in this session.
- 4. To submit your criteria and create a Hierarchical Clustering Analysis report, click the **Submit** button.

Performing Correlation Scatter Plot Analysis

A *Correlation Scatter Plot analysis* enables you to select two continuous variables and plot them against each other. The variables can be gene expression values or a clinical parameter like MRI percent longest diameter change.

The following are examples of how you can use a Correlation Scatter Plot Analysis.

- **Cross platform validation**: Select the same gene on two different platforms and display the correlation between the expression values.
- **Interreporter validation**: Select the same gene (but different reporters) on the same platform.
- **Gene expression correlation**: Investigate the relationships between gene expression values for two different genes.
- Clinical parameter and gene expression relationship: Investigate the relationship between a clinical parameter and the gene expression values for a given gene.

To perform a Correlation Scatter Plot Analysis, follow these steps:

1. The Correlation Scatter Analysis Form (*Figure 5.3*) enables you to fill in criteria for generate a correlation scatter plot.

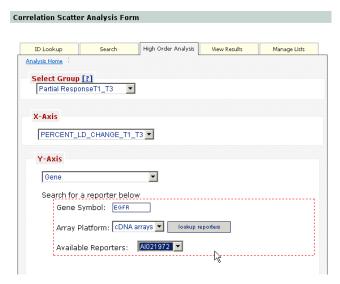


Figure 5.4 Selecting Correlation Scatter Plot criteria

2. You are required to enter at least one step for the correlation scatter plot. *Table 5.4* lists the available criteria:

Criteria	Special Instructions	
Select Group	A <i>group</i> is a pre-defined or user-defined PatientDID list comprising patient identifiers with certain characteristics. Select a group to filter the query to a collection of patients. Lists that you created appear in red.	
X-Axis	Enter gene information or select an LD_CHANGE option to show the correlation between the X- and Y-axis items.	
• Gene	Select Gene. Note that to generate the plot, you must select a gene symbol for one axis. • Enter a gene symbol. • Select an array. • Click the Lookup Properties button. • Select a reporter.	

Table 5.4 Correlation Scatter Plot criteria instructions

Criteria	Special Instructions
PERCENT LD_CHANGE	This group of options enables you to analyze the percentage of LD (Longest Diameter) change in the size of the tumor. • PERCENT_LD_CHANGE: On the X-axis include all samples and compare against all PERCENT_LD_CHANGE values. • PERCENT_LD_CHANGE_T1_T2: On the X-axis display the percentage of LD change between timepoints T1 and T2.
	 PERCENT_LD_CHANGE_T1_T3: On the X-axis display the percentage of LD change between timepoints T1 and T3. PERCENT_LD_CHANGE_T1_T4: On the X-axis display the percentage of LD change between timepoints T1 and T4.
Y-Axis	Same as the X-axis options.
Correlation	 Pearson correlation measures the relative shape of the gene regulations rather than the absolute levels. This is a natural choice, because it is widely used to measure gene correlations. Spearman correlation is a non-parametric test for the strength of the relationship between pairs of variables. Spearman Rank Correlation measures the correlation between two sequences of values. The two sequences are ranked separately and the differences in rank are calculated at each position. The range of Spearman Correlation is from -1 to 1. Spearman Correlation can detect certain linear and non-linear correlations. However, Pearson Correlation may be more appropriate for finding linear correlations.

Table 5.4 Correlation Scatter Plot criteria instructions

- 3. You must enter a title/name for this analysis in the **Name Analysis Result** text box. This name must be unique among all your queries in this session.
- 4. To submit your criteria and create a Correlation Scatter Plot report, click the **Submit** button.

Performing Categorical Plot Analysis

The Categorical Plot analysis enables you to select one or more groups of patients as defined in the I-SPY Manage Lists function, and view a box-and-whisker plot of a continuous variable for the patients in the selected groups. These groups can be the pre-defined groups defined in I-SPY or groups that you create with the Manage Lists function. The following example describes how creating lists in the I-SPY Manage Lists function (see Combining Existing Lists to Create a New List on page 47) can generate categorical plots for specific needs.

- Using the I-SPY Manage Lists function, create two lists
 - A Triple Negative list combining the ER-, HER2-, PR- groups
 - A Triple Positive list combining the ER+, HER2+, PR+ groups
- Specify Categorical Plot criteria to compare the Percent Longest Diameter Change for timepoints T1 to T4 with values for patients in the Triple Positive group versus patients in the Triple Negative group.

Other general uses of box-and-whisker plots include the following:

- Indicate whether a distribution is skewed and whether there are potential unusual observations (outliers) in the dataset.
- Perform a large number of observations.
- Compare two or more datasets.
- Compare distributions because the center, spread, and overall range are immediately apparent.

To perform a Categorical Plot Analysis, follow these steps:

1. The Categorical Plot Analysis Form (*Figure 5.3*) enables you to fill in the criteria to generate a categorical plot.

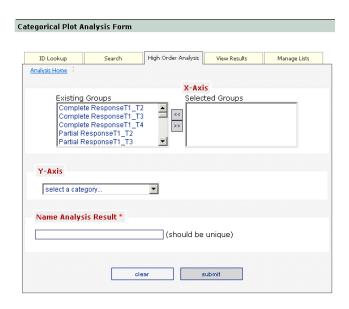


Figure 5.5 Selecting Categorical Plot Analysis criteria

2. You are required to complete an entry for the categorical plot. *Table 5.5* lists the available criteria:

Criteria	Special Instructions
Select Group (X-axis)	A <i>group</i> is a pre-defined or user-defined PatientDID list comprising patient identifiers with certain characteristics. Select the groups for the X-axis to filter the query to a collection of patients. Lists that you created appear in red.
Y-axis	Enter gene information or select an LD_CHANGE option to show the correlation between the X- and Y-axis items.
• Gene	 Select Gene for the X-Axis. Enter a gene symbol. Select an array. Click the Lookup Properties button. Select a reporter.
PERCENT LD_CHANGE	 This group of options enables you to analyze the percentage of LD (Longest Diameter) change in the size of the tumor. PERCENT_LD_CHANGE: On the X-axis include all samples and compare against all PERCENT_LD_CHANGE values. PERCENT_LD_CHANGE_T1_T2: On the X-axis display the percentage of LD change between timepoints T1 and T2. PERCENT_LD_CHANGE_T1_T3: On the X-axis display the percentage of LD change between timepoints T1 and T3. PERCENT_LD_CHANGE_T1_T4: On the X-axis display the percentage of LD change between timepoints T1 and T3.

Table 5.5 Correlation Plot Analysis criteria instructions

- 3. You must enter a title/name for this analysis in the **Name Analysis Result** text box. This name must be unique among all your queries in this session.
- 4. To submit your criteria and create a report, click the **Submit** button.

CHAPTER 6 VIEWING RESULTS

This chapter describes reports and search results that I-SPY returns after queries and and high order analyses.

Topics in this chapter include the following:

- Results Overview on page 33
- Search Results on page 33
- High Order Analysis Results on page 37

Results Overview

The View Results page shows a collection of reports previously viewed in a particular user session. This allows you to compare reports by opening them in separate windows. You can view results generated with the Search function and the High Order Analysis function.

Search Results

The following results are generated from the Search function:

- Clinical Reports
- IHC Level of Expression Search Results
- IHC Loss of Expression Search Results

View Results (*Figure 6.1*) displays the query name and lists the output generated for the query.

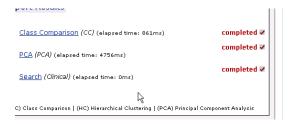


Figure 6.1 Search Results

Clinical Reports

A *Clinical report* displays the demographic, clinical, MR, and pathology data for a given set of patients (*Figure 6.2*). From any Clinical report in I-SPY, you can also create a PatientDID list to further filter a query or perform analysis. For more information, see the Related Topics.

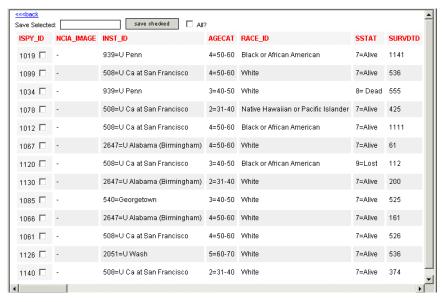


Figure 6.2 Clinical Report

Creating a PatientDID List from a Clinical Report

On any Clinical page, you can select and save patients to a PatientDID list.

- 1. There are two ways to select patients on the Clinical window:
 - To select an individual, select the box in the **I-SPY ID** column (*Figure 6.3*).



Figure 6.3 Checking the I-SPY ID column

To select all of the patients, select the All box (Figure 6.4).



Figure 6.4 Selecting all of the samples on the Clinical window

To clear all of the patients, uncheck the All box.

2. To save the patients to a PatientDID list, enter a name for the list (*Figure 6.5*).



Figure 6.5 Saving Selected Samples on the Clinical page

- 3. Click the **save checked** button. Sample List Saved appears.
- 4. Click the **OK** button.

The new PatientDID list is now displayed in red in the blue panel at the bottom of the PatientDID names (*Figure 6.6*). Mouse over the name and the data items appear.



Figure 6.6 Saved PatientDID list

IHC Level of Expression Search Results

An *IHC Level of Expression report* displays the patients selected from the IHC Level of Expression search (*Figure 6.3*). On the IHC Level of Expression page, you can select and save patients to a PatientDID list.

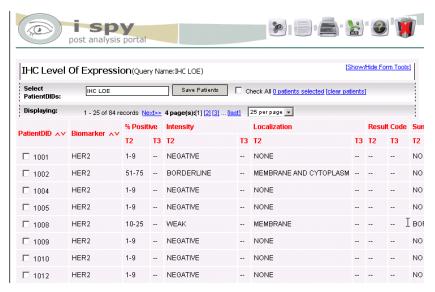


Figure 6.7 IHC Level of Expression Report

IHC Loss of Expression Search Results

An *IHC Loss of Expression report* displays the patients selected from the IHC Loss of Expression search (*Figure 6.3*). On the IHC Level of Expression page, you can select and save patients to a PatientDID list.

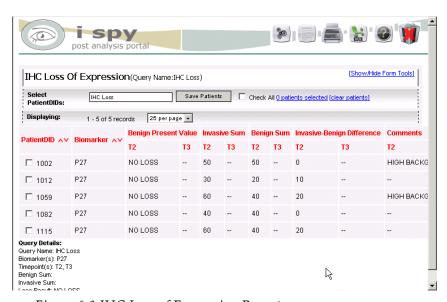


Figure 6.8 IHC Loss of Expression Report

High Order Analysis Results

The following reports are generated from the High Order Analysis function:

- Class Comparison Report
- Principal Component Analysis Plot
- Hierarchical Clustering Report
- Correlation Scatter Plot
- Categorical Plot Analysis

View Results (*Figure 6.9*) displays the query name and lists the output generated for the query.

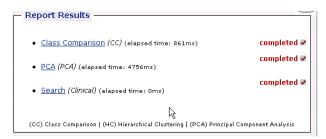


Figure 6.9 HOA Results

Class Comparison Report

The Class Comparison report (Figure 6.10) displays group average, fold change, and p-value based on the search parameters that you selected. For a **T-test** or **Wilcoxon** Statistical Method analysis (Figure 6.10), the Class Comparison report is as follows.

- The report displays the group average, where the numerator is the mean of log(base 2) expression signals from the samples in the first group. The denominator is the mean of log(base 2) expression signals from the samples in the second group.
- The fold change for the reporter between the selected groups appears along with *p*-value.
- Gene symbol annotations appear for each reporter. To obtain extensive annotations, click the Excel icon on the upper right-hand corner of the report.



Figure 6.10 Class Comparison Report

Creating a Gene List (Select Genes toolbar)

On the Class Comparison page, you can select and save the genes to a Gene list. You can use a Gene list to filter a query or perform analysis. To create a Gene list, follow these steps (*Figure 6.11*):

- 1. To select all of the genes in result list, click the Check All box.
- 2. To select some of the genes, check the box in the Gene Symbol column.

Note: To clear the selected genes, click the **clear genes** link.

 To save the selected genes, enter a unique name for the file next to Select Genes, or maintain the current name, which varies based on the type of Statistical Method selected for the analysis.



Figure 6.11 Selecting Genes instructions

4. Click the Save Genes button.

The results are saved.

5. Click the **OK** button.

The new Gene list appears in red in the blue panel at the bottom of the Gene Lists names. Mouse over the name and the data items appear.

Resorting Column Results

To sort a column in a report, follow these steps:

1. If a report column has red triangles pointing up and down next to the name, you can sort a column of numeric or alphabetical values (*Figure 6.12*).

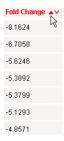


Figure 6.12 Sorting column results

2. To sort a column in ascending order, select the red triangle pointing up. To sort a column in descending order, select the red triangle pointing down.

Showing Additional Information

When results are listed in a report, row or column items may appear as links. Click the link to display additional information about the item.

For example, to display more information about a gene, click the name link (*Figure 6.13*).



Figure 6.13 The Gene column

The Cancer Genome Anatomy Project (CGAP) browser opens.

Principal Component Analysis Plot

The *Principal Component Analysis plot* (*Figure 6.14*) is a two-dimensional graph which plots the various principal components from the analyses.

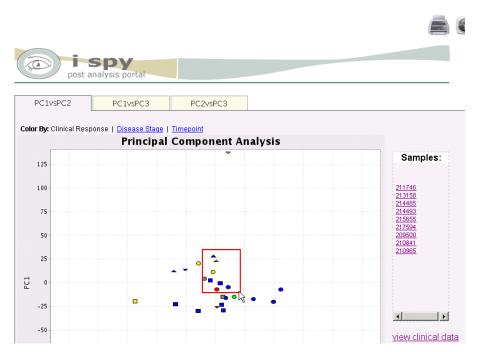


Figure 6.14 Principal Component Analysis Plot

Table 6.1 describes other areas in the plot:

Area	Description
Tabs	You can click on the three tabs at the top of the graph to display the following:
	PC1vsPC2 Pod Pod
	PC1vsPC3PC2vsPC3
Color By	Each point on the graph represents a sample, and by default, the samples are colored by Clinical Response . To color by Disease stage or Timepoint , click the appropriate link.
Legend	At the bottom of the graph, a legend defines how the different shapes on the graph indicate different survival ranges for patients.
Samples	The Samples area enables you to select, review, and display clinical data for samples in the plot (see <i>Selecting Samples of Interest in a Plot</i>).

Table 6.1 Areas of the Principal Component Analysis Plot

Selecting Samples of Interest in a Plot

To select the samples of interest in an I-SPY plot, follow these steps:

1. Click and drag a rectangle around the samples.

A red rectangle appears around the samples, and the list of the samples appears on the right-hand side (*Figure 6.15*).

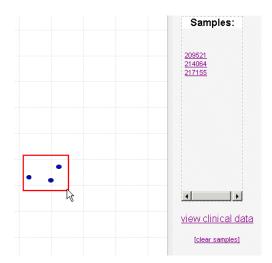


Figure 6.15 Sampling from a clinical plot

2. To help lasso the points on the plot and identify the location of these points, mouse over a sample name in the list.

A yellow circle appears on the plot where the sample is located (Figure 6.16).

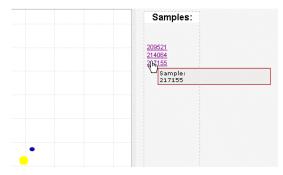


Figure 6.16 Lasso the points

3. To generate clinical data for the selected samples and save the samples, click the **view clinical data** link.

To select another group of samples, click the **clear samples** link and start again.

Hierarchical Clustering Report

The *Hierarchical Clustering report* (*Figure 6.17*) displays the dendrogram from the hierarchical clustering analysis and a clinical report. The dendrogram is organized based on the gene expression profiles of the samples. Samples with similar profiles are placed closer together on the tree. To adjust the size of the graph, move the box on the **Image Control** bar in the top lefthand corner.

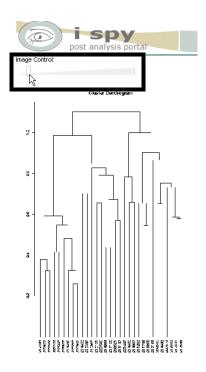


Figure 6.17 Hierarchical Clustering Dendrogram

A clinical report appears beneath the dendrogram where you can select patients and create a PatientDID list. You can also display an image associated with a patient. Click the NCIA icon in the **NCIA_Image** column (*Figure 6.18*), and the National Cancer Imaging Archive(NCIA) web site appears. As a first time user, register to obtain a username and password, and then you will have access to these images.



Figure 6.18 Hierarchical Clustering Clinical Report

Correlation Scatter Plot

The Correlation Scatter plot (Figure 6.19) is a visualization used to compare two continuous variables. The X-axis represents the values for one of the variables and the Y-axis represents the values for the other variable.

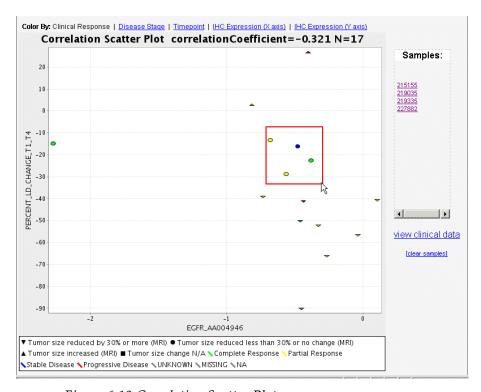


Figure 6.19 Correlation Scatter Plot

Table 6.1 describes other areas in the plot:

Area	Description
Color By	Each point on the graph represents a sample, and by default, the samples are colored by Clinical Response . To color by Disease stage or Timepoint , click the appropriate link.
	To plot IHC expression on the X- or Y-axis, click the IHC Expression X axis or IHC Expression Y axis link. For example, selecting color by IHC Expression (X Axis), colors the points on the plot based on the IHC Expression of the gene on the X-axis. Selecting color by IHC Expression (Y Axis) colors the points on the plot based on the IHC Expression of the gene on the Y-axis.
Correlation Coefficient	Computed and displayed in the title. Correlation coefficients with values close to 1 are highly correlated. Values close to -1 indicate an inverse relationship. Values close to 0 indicate no correlation between the parameters.

Table 6.2 Areas of the Correlation Scatter Plot

Area	Description
Legend	At the bottom of the graph, a legend defines how the different shapes on the graph indicate different survival ranges for patients.
Samples	The Samples area enables you to select, review, and display clinical data for samples in the plot (see <i>Selecting Samples of Interest in a Plot</i>).

Table 6.2 Areas of the Correlation Scatter Plot

Categorical Plot Analysis

The Categorical Plot analysis (Figure 6.20) displays a box-and-whisker plot of a continuous variable for patients in selected groups.

The following items in the graph indicate the following:

- Black dot in the box indicates mean value.
- Horizontal line in the box indicates the median value.
- Circles are potential outliers.
- Triangles are outliers beyond the graph.

Example uses of box-and-whisker plots include the following:

- Indicate whether a distribution is skewed and whether there are potential unusual observations (outliers) in the dataset.
- Perform a large number of observations
- Compare two or more datasets.
- Compare distributions because the centre, spread, and overall range are immediately apparent.

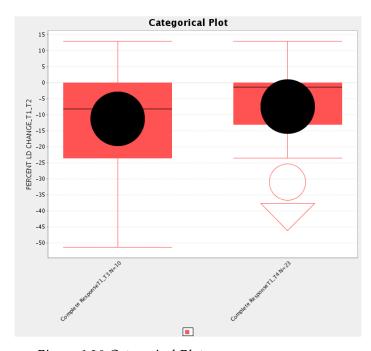


Figure 6.20 Categorical Plot

CHAPTER 7 MANAGING LISTS

This chapter describes how to manage patient and gene lists.

Topics in this chapter include:

- Managing Lists Overview on page 45
- Viewing the Data Items in a List on page 46
- Removing Data Items to Create a New List on page 47
- Deleting an Entire List on page 48
- Adding a New "Custom" List on page 49
- Combining Existing Lists to Create a New List on page 47

Managing Lists Overview

The I-SPY Manage Lists function centralizes all activities pertaining to the creation and management of user-defined, as well as study-defined **PatientDID Lists** and **Gene Lists**. With these lists, you can further refine queries or facilitate analysis. Using the Manage List function, you can perform the following functions:

- View data items in a list
- Create new lists from existing lists
- Delete lists
- Add lists by uploading them or typing them

PatientDID Lists

An I-SPY *PatientDID list* is a list of patients with certain characteristics that you can use to filter a query or perform analysis. These lists are pre-defined in the Manage Lists function, or you can create PatientDID lists from the ID Lookup function, IHC Search results, and Hierarchical Clustering results. The blue panel displays all PatientDID lists. Mouse over a list name in the blue panel to display the list's data items. Note that user-defined lists appear in red type in the blue panel.

Gene Lists

An I-SPY *Gene list* is a list of genes with certain characteristics that you can use to filter a query or perform analysis. The pre-defined gene list is **defaultGene1**, or you can create your own list with the Class Comparison results. The blue panel displays the names of the Gene lists. Mouse over a list name in the blue panel to display the list's data items. Note that user-defined lists appear in red type in the blue panel.

Viewing the Data Items in a List

To view the individual data items on a list, follow these steps:

1. At the top of the Manage List page, click on the type of lists you would like to view (PatientDID List or Gene List).

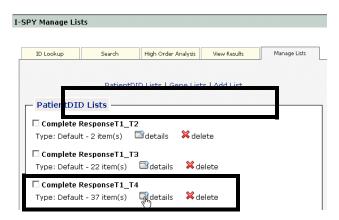


Figure 7.1 List types and Details

Find a list to be viewed, and click the **details** icon to display all of the items in the list.

Note: You can also mouse-over the list name in the blue panel and the list's data item names appear.

3. To export the list, click the **export list** link at the bottom of the data item list.

Removing Data Items to Create a New List

You may delete items from an existing list, then view the new list or save the list. To delete data items, follow these steps:

- 1. At the top of the Manage List page, click on the type of lists you would like to view (**PatientDID List** or **Gene List**).
- 2. Find the list you want to change, and click on the box next to the list name.
- 3. Click the **details** icon to display all the items in the selected list.

Figure 7.2 Deleting data items

4. Click the **delete** link beside the item you want to delete. The item is removed from the list.

Once you remove the items, you can view the new list or save the list to your computer.

- 5. Click the **export link** at the bottom of the items list.
- 6. Click **Open** or **Save**.

Combining Existing Lists to Create a New List

You may create new lists from existing lists. To create a custom list from existing lists, follow these steps:

- 1. At the top of the Manage List page, click on the type of lists you would like to view (PatientDID List or Gene List).
- 2. Click the box next to the list(s) to be used for creating a new list.

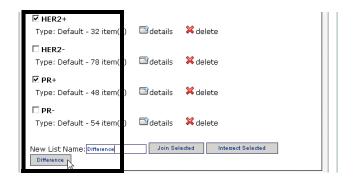


Figure 7.3 Combining existing lists

3. In the **New List Name** box, enter a unique name for the new list you are creating, and then click the appropriate button:

- Join combines two or more categories into a on list.
- Intersect creates a new list from only the items that appear on more than one selected list category.
- Difference creates up to two lists each comprising items that appeared in one of the selected lists. For example, if you select HER2+ and PR+, the new lists are "HER2+ - PR+" comprising the items that appeared in the HER2 list only and "PR+ - HER2+" comprising the items appearing in the PR+ list only (Figure 7.4).



Figure 7.4 New Difference lists

The new list(s) appear on the Manage Lists page and in the blue panel in red (Figure 7.4).

Deleting an Entire List

To delete one or more lists from a list type, follow these steps:

- 1. At the top of the Manage List page, click on the type of lists you would like to view (**PatientDID List** or **Gene List**).
- 2. Find the list you want to delete, and click the box next to the list name. Click more than one box to select multiple lists for deletion.



Figure 7.5 Deleting an entire list

3. To delete the selected lists, click an **x delete** icon. The selected categories are removed.

Adding a New "Custom" List

You may add a new list type by *uploading* a list from your computer or *manually creating* a list. To add a new list, follow these steps:

- 1. At the top of the Manage List page, click Add List.
 - The **Upload List or Manually type List** block appears.
- 2. To upload a list, follow these steps:
 - 1. Click **Upload List** at the top of the box.



Figure 7.6 Uploading a list

- 2. From the **Choose the list type** drop-down list box, select the list to be uploaded.
- 3. Click the **Browse** button beside the **Upload file** box. Navigate to and select the file to be uploaded.

Note: The list must have one ID on each line, and the file must end with a Return.

- 4. Enter a unique name for the list, and then click the **Add List** button. The new list appears on the blue panel in red.
- 3. To create and add a list manually, follow these steps:
 - 1. Click **Manually Type List** at the top of the box.

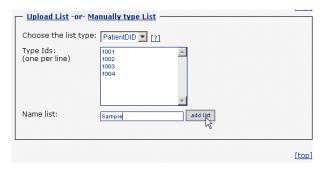


Figure 7.7 Manually typing a list

- 2. From the **Choose the list type** drop-down list box, select the list to be uploaded.
- 3. In the **Type Ids** box, enter items into the text block by typing them one to a line. End the list with a Return.

- 4. Enter a unique name for the list, and then click the **Add List** button. The new list appears on the blue panel in red.
- 5. To open and view the newly created list or save it to your computer, click on the list name in the blue panel. Click **Open** or **Save**.
- 6. To open and view the newly created list or save it to your computer, click on the list name in the blue panel. Click **Open** or **Save**.

GLOSSARY

Acronyms and other terms referred to in the chapters of this User's Guide are described in this glossary.

Term	Definition
NCIA	National Cancer Imaging Archive
Class Comparison	Differential gene expression across the tumor types will be evaluated by calculating the typical <i>t</i> -statistic for each reporter. Both parametric and non-parametric <i>p</i> -value will be computed.
False Discovery Rate (FDR)	The expected proportion of Type I errors among rejected hypotheses in simultaneous testing of multiple null hypotheses.
Family-wise Error Rate (FWER)	Denotes the probability of having at least one false significant test result within the set of tested hypotheses.
Gene List	A pre-defined or user-defined list in I-SPY comprising genes with a set of characteristics. Used to filter a query.
Hierarchical Clustering	Hierarchical cluster analysis is a statistical method for finding relatively homogeneous clusters of cases based on measured characteristics. It starts with each case in a separate cluster and then combines the clusters sequentially, reducing the number of clusters at each step until only one cluster is left.
High Order Analysis	After data preprocessing (filtering and normalization), further statistical analysis of gene expression data are performed, including class comparison, class discovery and class prediction.

Table 8.1 Glossary of I-SPY terms

Term	Definition
IHC (Immunohistochemistry)	Method of analyzing and identifying cell types based on the binding of antibodies to specific components of the cell. It is sometimes referred to as immunocytochemistry.
NCI	National Cancer Institute
NCICB	National Cancer Institute Center for Bioinformatics
PatientDID List	A pre-defined or user-defined list in I-SPY comprising patients with a set of characteristics. Used to filter a query.
Principal Component Analysis	PCA is commonly used in microarray research as a tool. It is designed to capture the variance in a dataset in terms of principle components. In effect, one is trying to reduce the dimensionality of the data to summarize the most important (i.e. defining) parts while simultaneously filtering out noise.
Wilcoxon Test	Nonparametric statistics for testing hypotheses about whether two samples differ.

Table 8.1 Glossary of I-SPY terms

INDEX

Symbols 38	black dot in box defined 44 circles defined 44 High Order Analysis 30
A	line in box defined 44 triangles defined 44
Across Timepoints, definition 22	uses for 44
Adding	uses for Box and Whisker plot 30
custom list (uploading or typing) 49	using Managing Lists 30
new list (combining) 47	Circles
new list (removing data items) 47	on Categorical Plot 44
user account 4	Class Comparison
Analyses	Analysis, definition 22
Categorical Plot 30	Analysis Form, description 22
Class Comparison 22	example (View Results) 37
Correlation Scatter Plot 28	High Order Analysis 22
Hierarchical Clustering 26	report description (View Results) 37
Principal Component 24	Clinical plot
Audience	selecting samples of interest 41
of the User's Guide v	Clinical Query
Average linkage, definition 27	definition 13
_	filling in form 14
В	Form page, description 14
Baseline, assigning 23	Clinical report
Black dot in box	creating PatientDID list 35
on Categorical Plot 44	description 34
Blue panel	saving genes 38
description 6	saving samples 35
red Gene list (Class Comparison) 38	selecting samples 35
red name user created 35	Combining lists 47
red PatientDID list (Clinical report) 35	Complete linkage, definition 27
red PatientDID list (Search) 12	Constrain by GeneList, definition 25, 26
blue panel	Constrain reporters by variance (Gene Vector)
manage list information 46	percentile, definition 25, 26
Box and Whisker plot	Correlation, definition 29
Categorical Plot 30	Correlation Coefficient, definition 43
C	Correlation Scatter Plot
C	Analysis, definition 27
Catagorical Plot	Analysis Form, description 28
Categorical Plot	examples of 27
Analysis, description 30	High Order Analysis 28
Analysis, report, description 44	report, description 43
Analysis Form, description 30	Creating

custom list (uploading or typing) 49	Correlation Scatter Plot 28
new list (combining) 47	Hierarchical Clustering 26
new list (removing data items) 47	overview 21
user account 4	Principal Component 24
	Horizontal line in box
D	on Categorical Plot 44
Data items	_
removing from a list 47	I
viewing in a list 46	ID Lookup function
defaultGene1, description 26	creating PatientDID list 12
Deleting	displaying patient sample information 11
list 48	downloading data 12
Difference option, description 48	overview 9
Distance Matrix, definition 27	steps 10
	IHC Expression X/Y axis, Color By, definition 43
Downloading	IHC Level of Expression
ID Lookup data 12	defined 16
T.	filling in Query form 16
E	report, description 36
E-mailing	report, example 36
NCICB Application Support 4	IHC Loss of Expression
Euclidean distance, definition 27	defined 18
Excel file	
downloading ID Lookup data 12	filling in Query form 18 Form page, description 18
8	1 0 1
F	report, example 36
	report, example 36
Fixed Timepoint, definition 22	Intersect option, description 48
Fold Change, definition 23	I-SPY
	creating user account 4
G	description 1
Gene list	functions 2
creating with Manage Lists 47	getting help 7
from Manage Lists 25	launching 3
saving on Class Comparison report 38	logging in 5
selecting and saving 38	logging out 8
Group	menu 7
selecting 14	workspace description 6
	workspace screen 6 ISPY ID
H	definition 11
	definition 11
Help	T
obtaining 7	J
Help link 7	Join option, description 48
Hierarchical Clustering	
Analysis, definition 26	L
Analysis Form, description 26	LabTrak ID
dendogram description 42	definition 11
High Order Analysis 26	Launching
NCIA_Image icon description 42	I-SPY 3
page description (View Results) 42	
report, description 42	Legal Rules of the Road page 5
High Order Analysis function	Line in box
Categorical Plot 30	on Categorical Plot 44
Class Comparison 22	Linkage Method, definition 27

Lists	report example 40
adding a custom list 49	Purpose
combining lists 47	of the User's Guide v
creating a new list 47	p-value, definition 23
deleting 48	_
removing data items 47	R
viewing data items 46	Red list name
Logging in	
I-SPY 5	user created list 35
Logging out	Removing
I-SPY 8	data items in a list 47
Logout link 8	Reports
O	Categorical Plot 44
M	Class Comparison 37
	Correlation Scatter Plot report example 43
Managing Lists function	Hierarchical Clustering 42
creating PatientDID list 35	Principal Component Analysis 40
for Categorical Plot analysis 30	viewing 33
overview 45	
Manually Type List options, description 49	S
Microsoft Excel	Sample
downloading ID Lookup data 12	selecting on Clinical report 35
Morphology, definition 15	selecting on plot 41
	Sample ID
N	displaying patient information 11
NCIA, definition 42	searching for 10
, , , , , , , , , , , , , , , , , , , ,	Samples area
0	selecting samples 41
	Samples of interest
Overview	41
High Order Analysis function 21	Saving
ID Lookup function 9	genes on Class Comparison 38
managing lists 45	samples on Clinical report 35
Search function 13	Search function
View Results function 33	Clinical Query 14
D.	IHC Level of Expression Query 16
P	IHC Loss of Expression Query 18
PatientDID list	overview 13
creating with ID Lookup 12	Searching
creating with Manage Lists 47	patient IDs and samples 10
description 46	Select Genes toolbar 38
in Clinical report 35	Select Group
saving on Clinical report 35	assigning 14
user list blue panel 35	Single linkage, definition 27
Patient ID	Sorting
displaying sample information 11	O
searching for 10	report columns 39
Pearson correlation, definition 27, 29	Spearman, definition 29
Percent LD, definition 15	Statistical Method, definition 23
PERCENT LD_CHANGE, definition 29, 31	Support link 7
Principal Component	
Analysis, defined 24	T
Analysis Form, description 24	Timepoint, definition 11
Analysis report, description 40	Triangles
High Order Analysis 24	on Categorical Plot 44

```
top of columns to sort 39
T-test Two Sample Test
   definition 23
   output 37
Typing
   a list (Manage List function) 49
U
Uploading
   list (Manage List function) 49
User's Guide
   audience v
   organization of v
   purpose v
   text conventions vi
User account
   creating 4
User guide link 7
\mathbf{V}
Viewing
   data items in a list 46
View Results function
   overview 33
W
Wilcoxon Test Man-Whitney Test
   definition 23
   output 37
Y
Yellow circle
   meaning on plot 41
```