

I-SPY USER'S GUIDE

Version 1.0



NATIONAL[®]
CANCER
INSTITUTE

Center for Bioinformatics

June 21, 2007

CREDITS AND RESOURCES

<i>I-SPY</i> <i>Development and Management Teams</i>		
<i>Development</i>	<i>Documentation</i>	<i>Product and Program Management</i>
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>I-SPY Development and Management Teams</i>		
<i>Development</i>	<i>Documentation</i>	<i>Product and Program Management</i>
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.		

<i>Contacts and Support</i>	
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	v
Topics Covered	v
Text Conventions Used	vi
 Chapter 1	
About I-SPY 1.0	1
About I-SPY	1
About I-SPY Functions	2
 Chapter 2	
Getting Started with I-SPY 1.0	3
Launching I-SPY	3
Creating a User Account	4
Logging In	5
Accepting I-SPY Provisions	5
Welcome to I-SPY 1.0	6
Getting Help	7
Application Support	7
Logging Out	8
 Chapter 3	
Conducting an Identifier Lookup	9
ID Lookup Overview	9
Looking Up a Patient or Sample Identifier	10
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 Searches13

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 Analysis21

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

Chapter 6

Viewing Results33

Results Overview	33
Using I-SPY Report Icons	34
Search Results	34
Clinical Reports	35
IHC Level of Expression Search Results	37
IHC Loss of Expression Search Results	37
High Order Analysis Results	38
Class Comparison Report	38
Principal Component Analysis Plot	40
Hierarchical Clustering Report	42
Correlation Scatter Plot	43
Categorical Plot Analysis	44

Chapter 7

Managing Lists47

Managing Lists Overview	47
PatientDID Lists	48
Gene Lists	48
Viewing the Data Items in a List	48
Removing Data Items to Create a New List	49
Combining Existing Lists to Create a New List	49
Deleting an Entire List	50
Adding a New "Custom" List	51

Data Dictionary	53
I-SPY Data Dictionary	54
Patient Demographic Data Dictionary	54
Chemotherapy Summary Data Dictionary	56
On-Study Data Dictionary	57
Post-Surgery Summary Data Dictionary	58
Follow-Up Data Dictionary	60
Response Evaluation Data Dictionary	60
Glossary	63

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 identifiers. 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.

- [Appendix A](#) provides a data dictionary listing the variables for I-SPY.
- [Glossary](#) describes terms used in this guide.

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 .
<u>URL</u>	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.
<i>Italics</i>	Highlights references to other documents, sections, figures, and tables.	See <i>Figure 4.5</i> .
Italic boldface monospace type	Represents text that you type.	In the New Subset text box, enter <i>Proprietary Proteins</i> .
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 1 ABOUT I-SPY

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 multi-center 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. See Conducting an Identifier Lookup .
Perform a Search	Perform one of the following types of queries: <ul style="list-style-type: none">• Clinical query• IHC Level of Expression query• IHC Loss of Expression query See Conducting Searches .
Perform a High Order Analysis	Run the following types of higher order analyses : <ul style="list-style-type: none">• Class comparisons• Hierarchical clustering• Principal component analyses• Correlation scatter plot• Categorical plot analysis See High Order Analysis .
View Results	View Search and High Order Analysis results. See Viewing Results .
Manage Lists	Manage user-defined or study-defined patient or gene identifier lists. You can use them to filter queries or perform analysis. See Managing Lists .

Table 1.1 I-SPY user tasks

GETTING STARTED WITH I-SPY

This chapter introduces general I-SPY procedures and how to obtain help to use I-SPY.

Topics in this chapter include:

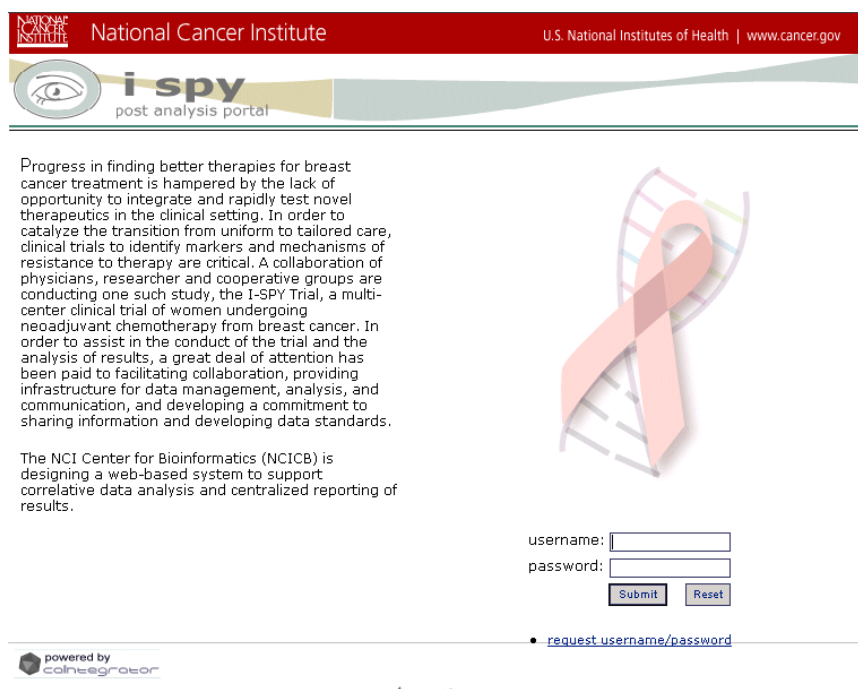
- *Launching I-SPY* on page 3
- *Creating a User Account* on page 4
- *Logging In* on page 5
- *Accepting I-SPY Provisions* on page 5
- *Welcome to I-SPY 1.0* on page 6
- *Getting Help* on page 7
- *Application Support* on page 7
- *Logging Out* on page 8

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*).



National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov

i spy
post analysis portal

Progress in finding better therapies for breast cancer treatment is hampered by the lack of opportunity to integrate and rapidly test novel therapeutics in the clinical setting. In order to catalyze the transition from uniform to tailored care, clinical trials to identify markers and mechanisms of resistance to therapy are critical. A collaboration of physicians, researcher and cooperative groups are conducting one such study, the I-SPY Trial, a multi-center clinical trial of women undergoing neoadjuvant chemotherapy from breast cancer. In order to assist in the conduct of the trial and the analysis of results, a great deal of attention has been paid to facilitating collaboration, providing infrastructure for data management, analysis, and communication, and developing a commitment to sharing information and developing data standards.

The NCI Center for Bioinformatics (NCICB) is designing a web-based system to support correlative data analysis and centralized reporting of results.

username:
password:

[request username/password](#)

powered by
colintegrator

ALTERNATIVE

Figure 2.1 I-SPY login page

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**.



username:

password:

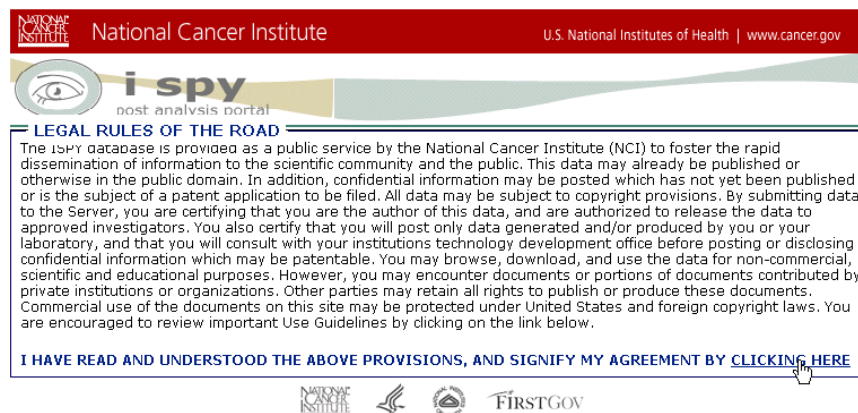
• [request username/password](#)

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.



National Cancer Institute U.S. National Institutes of Health | www.cancer.gov

i-spy
post analysis portal

LEGAL RULES OF THE ROAD

The I-SPY database is provided as a public service by the National Cancer Institute (NCI) to foster the rapid dissemination of information to the scientific community and the public. This data may already be published or otherwise in the public domain. In addition, confidential information may be posted which has not yet been published or is the subject of a patent application to be filed. All data may be subject to copyright provisions. By submitting data to the Server, you are certifying that you are the author of this data, and are authorized to release the data to approved investigators. You also certify that you will post only data generated and/or produced by you or your laboratory, and that you will consult with your institutions technology development office before posting or disclosing confidential information which may be patentable. You may browse, download, and use the data for non-commercial, scientific and educational purposes. However, you may encounter documents or portions of documents contributed by private institutions or organizations. Other parties may retain all rights to publish or produce these documents. Commercial use of the documents on this site may be protected under United States and foreign copyright laws. You are encouraged to review important Use Guidelines by clicking on the link below.

I HAVE READ AND UNDERSTOOD THE ABOVE PROVISIONS, AND SIGNIFY MY AGREEMENT BY [CLICKING HERE](#)

NATIONAL CANCER INSTITUTE U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES FIRSTGOV

Figure 2.3 Legal Rules of the Road page

The I-SPY workspace appears (Figure 2.4).

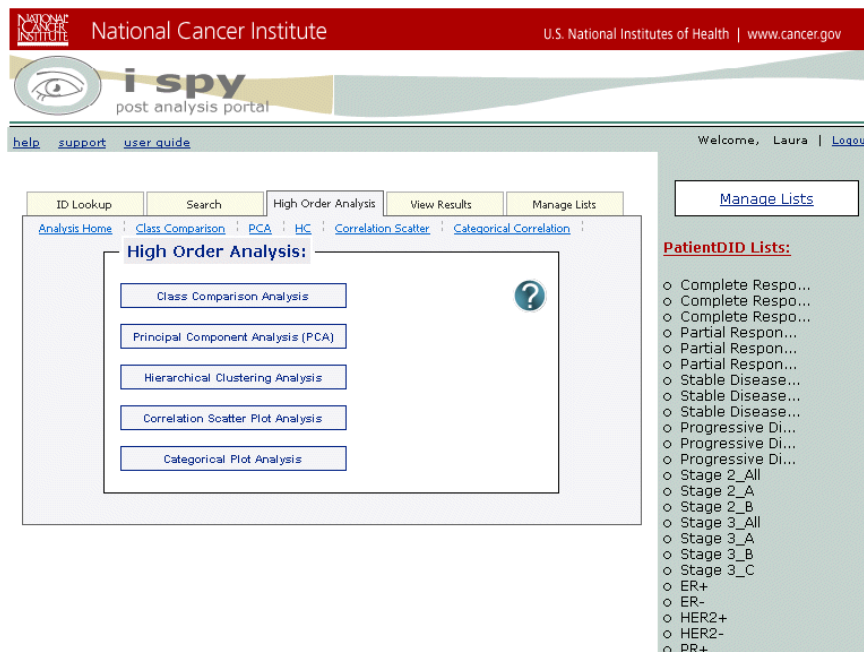


Figure 2.4 The I-SPY workspace

Welcome to I-SPY

The I-SPY workspace comprises a set of five tabs, a side, 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 side 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-SPY 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

The I-SPY login page appears.

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*

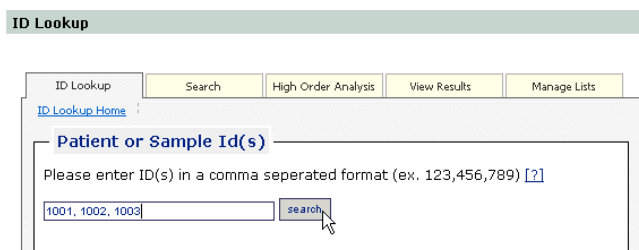
Note: You can use a [PatientDID list](#) to filter I-SPY queries and perform data analysis.

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.

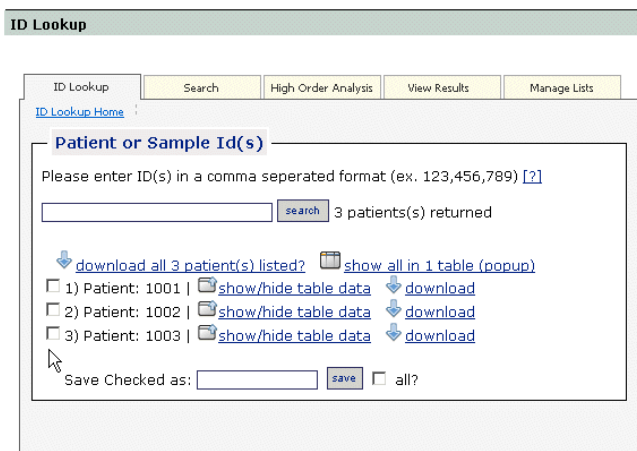


The screenshot shows the 'ID Lookup' page with a navigation bar containing 'Search', 'High Order Analysis', 'View Results', and 'Manage Lists'. Below the navigation bar is a search form titled 'Patient or Sample Id(s)'. The form contains a text input field with the text '1001, 1002, 1003' and a 'search' button. Above the input field is a prompt: 'Please enter ID(s) in a comma separated format (ex. 123,456,789) [?]'. There is also a link 'ID Lookup Home' above the form.

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*).




The screenshot shows the 'ID Lookup' page with the search results for the identifiers '1001, 1002, 1003'. The search button is now disabled, and the text '3 patients(s) returned' is displayed. Below the search button, there are three links: 'download all 3 patient(s) listed?', 'show all in 1 table (popup)', and a table of results. The table has three rows, each representing a patient. Each row contains a checkbox, the patient ID, a 'show/hide table data' link, and a 'download' link. At the bottom of the table, there is a 'Save Checked as:' field with a 'save' button and a checkbox labeled 'all?'.

checkbox	Patient ID	show/hide table data	download
<input type="checkbox"/>	1) Patient: 1001	show/hide table data	download
<input type="checkbox"/>	2) Patient: 1002	show/hide table data	download
<input type="checkbox"/>	3) Patient: 1003	show/hide table data	download

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  next to the patient's row.

The table highlights the lookup criteria (*Figure 3.3*).

[download all 3 patient\(s\) listed?](#)
[show all in 1 table \(popup\)](#)

☐ 1) Patient: 1001 |
 [show/hide table data](#)
[download](#)


	ISPY ID	LabTrak ID	Timepoint	Core Type	Section Info
1	1001	209512	T1	PARAFFIN	Pre-Tx FFPE Core 1
2	1001	209513	T1	PARAFFIN	Pre-Tx FFPE Core 2
3	1001	209514	T1	FROZEN	Pre-Tx Frozen Core 1
4	1001	209515	T1	FROZEN	Pre-Tx Frozen Core 2
5	1001	209516	T1	FROZEN	Pre-Tx Froz H&E Slide 1
6	1001	209517	T1	FROZEN	Pre-Tx Froz H&E Slide 2
7	1001	209518	T2	PARAFFIN	Anthr. C#1 FFPE Core 1


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	The timepoint when the sample was collected. <ul style="list-style-type: none"> • 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.

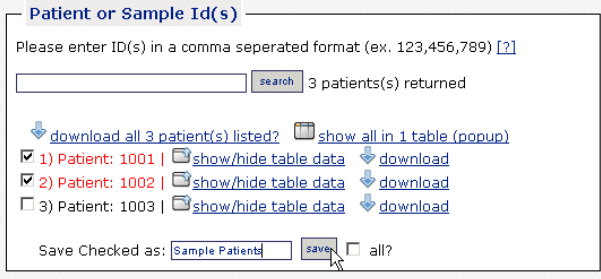
If you searched for *multiple* patients, to save all the patients' data to the same Excel file, follow these steps.

1. Click  above the list of patients to download all the patient sample information to one file.
 2. 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](#). 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.



Patient or Sample Id(s)

Please enter ID(s) in a comma separated format (ex. 123,456,789) [\[?\]](#)

3 patients(s) returned

[download all 3 patient\(s\) listed?](#) [show all in 1 table \(popup\)](#)

<input checked="" type="checkbox"/>	1) Patient: 1001	show/hide table data	download
<input checked="" type="checkbox"/>	2) Patient: 1002	show/hide table data	download
<input type="checkbox"/>	3) Patient: 1003	show/hide table data	download

Save Checked as: ☐ all?

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 side bar 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	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: <ul style="list-style-type: none"> • Stage 2_All • Stage 2_A • Stage 2_B • Stage 3_All • Stage 3_A • Stage 3_B • Inflammatory
	Agents	Select an agent to further filter the query.

Table 4.1 Clinical Query search criteria instructions

Criteria	Item Name	Special Instructions
	Clinical Response	<p>Select one or more clinical responses and the appropriate timepoint range.</p> <ul style="list-style-type: none"> • 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	<p>This group of options enables you to specify the percentage of LD (Longest Diameter) change in the size of the tumor between two timepoints.</p> <p>Select the timepoint range in which to analyze the percentage of LD change:</p> <ul style="list-style-type: none"> • 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. <p>Select the greater than/equal to (\geq) or less than/equal to (\leq) indicator and enter the percentage of LD change to search for in the selected timepoint range.</p>
Pathology	Pathology Tumor Size	Specify the tumor size and associated biomarkers to filter the query. Select the greater than/equal to (\geq) or less than/equal to (\leq) indicator and enter a value in centimeters of the tumor size.
	Status	<p>Specify the Pathology Status:</p> <p>ER+: Estrogen receptor positive ER-: Estrogen receptor negative PR+: Progesterone receptor positive PR-: Progesterone receptor positive HER2+: HER2 positive HER2-: HER2 negative</p>

Table 4.1 Clinical Query search criteria instructions

- 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 query 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 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*).

IHC Level of Expression Query Form

ID Lookup Search High Order Analysis View Results Manage Lists

[Search Home](#) [Clinical](#) [IHC Level](#) [IHC Loss](#) [FISH](#)

Select Group [?]

none

Select Timepoint *

Timepoint:

T1
T2
T3
T4

Select Biomarkers [?]

EGFR
FAK
HER2
KI-67

Intensity of Stain [?]

BORDERLINE
MODERATE/STRONG
NEGATIVE

Percent Positive [?]

between 0 % and 100 %

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	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: <ul style="list-style-type: none"> • P27 • Ki-67 • EGFR • CCND1 • P53 • HER2 • BCL2 • FAK
Intensity of Stain	Select an option that best describes the intensity of stain: <ul style="list-style-type: none"> • 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: <ul style="list-style-type: none"> • 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 stain: <ul style="list-style-type: none"> • None • Homogenous • Heterogenous

Table 4.2 IHC Level 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 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 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*).

IHC Loss of Expression Query Form

ID Lookup Search High Order Analysis View Results Manage Lists

Search Home > Clinical > IHC Level > IHC Loss > FISH

Select Group [?]
none

Select Timepoint *

Timepoint:
T1
T2
T3
T4

Select Biomarkers [?]
P27

Invasive Range [?]
invasive sum any 0

Benign Range [?]
benign sum any 0

Result Code [?]

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	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: <ul style="list-style-type: none"> • = • >= • <= Specify the value for the invasive sum.
Benign Range	Specify equal to, greater than, or less than to define the range benign sum: <ul style="list-style-type: none"> • = • >= • <= 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.
- 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* allows you to identify genes and reporters that are differentially expressed between two groups. To perform a [Class Comparison](#), follow these steps:

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

Class Comparison Analysis Form

Analysis Home

Select Timepoint* [?]

☒ Fixed Timepoint: T1 ☐ Across Timepoints: T1 vs. T2

Select Group* [?]

Existing Groups: Complete ResponseT1_T2, Complete ResponseT1_T3, Complete ResponseT1_T4, Partial ResponseT1_T2, Partial ResponseT1_T3

Selected Groups: (empty)

Baseline: none

Select Statistic

☒ Default ☐ Advanced

Select Array Platform

Agilent

Name Analysis Result *

Class Comparison (should be unique)

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	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 (analyzes one group at different timepoints).
	Baseline	For a <i>Fixed Timepoint</i> analysis, the baseline is determined by the second group in the Selected Groups box. For an <i>Across Timepoints</i> 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.
	<ul style="list-style-type: none"> Statistical Method 	Select the appropriate statistical method: <ul style="list-style-type: none"> 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.
	<ul style="list-style-type: none"> Multiple Comparison Adjustment 	Family-wise Error Rate (FWER) : Bonferroni False Discover Rate (FDR) : Benjamini-Hochberg
	<ul style="list-style-type: none"> 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.
	<ul style="list-style-type: none"> 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.

The screenshot shows the 'Principal Component Analysis (PCA) Form' page. At the top, there is a navigation bar with tabs: 'ID Lookup', 'Search', 'High Order Analysis', 'View Results', and 'Manage Lists'. Below the navigation bar, there is a link 'Analysis Home'. The main form area contains several sections: 'Select Timepoint *' with a dropdown menu showing 'T1', 'T2', 'T3', and 'T4'; 'Filter Genes/Reporters' with radio buttons for 'Default' and 'Advanced'; 'Select Array Platform' with a dropdown menu showing 'Agilent'; and 'Name Analysis Result *' with a text input field containing 'PCA Analysis' and a note '(should be unique)'. At the bottom of the form, there are two buttons: 'clear' and 'submit'.

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.
	<ul style="list-style-type: none"> 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.
	<ul style="list-style-type: none"> Constrain by GeneList 	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.
	<ul style="list-style-type: none"> 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.
	<ul style="list-style-type: none"> Constrain by GeneList 	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: <ul style="list-style-type: none"> • 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: <ul style="list-style-type: none"> • 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	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.
<ul style="list-style-type: none"> • Gene 	<p>Select Gene. Note that to generate the plot, you must select a gene symbol for one axis.</p> <ul style="list-style-type: none"> • 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
<ul style="list-style-type: none"> PERCENT_LD_CHANGE 	<p>This group of options enables you to analyze the percentage of LD (Longest Diameter) change in the size of the tumor.</p> <ul style="list-style-type: none"> 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	<p>Select a distance matrix option:</p> <p>Pearson correlation: Pearson's Correlation Coefficient measures the strength of the linear relationship between two variables. Assumptions are the following: linear relationship between two variables; continuous random variables; both variables must be normally distributed; and two variables must be independent of each other.</p> <p>Spearman correlation: Spearman's Rank Correlation Coefficient is a non-parametric measure of correlation. It assesses how well an arbitrary monotonic function could describe the relationship between two variables without making any assumptions about the frequency distribution of the variables. It does not require the assumption that the relationship between the variables is linear. It can be used for variables measured at the ordinal level.</p>

Table 5.4 Correlation Scatter Plot criteria instructions

- 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.
- 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 49) 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)	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.
<ul style="list-style-type: none"> Gene 	<p>Select Gene for the X-Axis.</p> <ul style="list-style-type: none"> Enter a gene symbol. Select an array. Click the Lookup Properties button. Select a reporter.
<ul style="list-style-type: none"> PERCENT LD_CHANGE 	<p>This group of options enables you to analyze the percentage of LD (Longest Diameter) change in the size of the tumor.</p> <ul style="list-style-type: none"> 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.

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
- [*Using I-SPY Report Icons*](#) on page 34
- [*Search Results*](#) on page 34
- [*High Order Analysis Results*](#) on page 38

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.

Using I-SPY Report Icons

Table 6.1 describes icons that appear at the top of most I-SPY report results:







Icon	Special Instructions
	Show or hide the report tools that appear below the report title and above the report data.
	Displays the query details at the bottom of the report used to generate the data.
	Prints the report.
	Saves the report to a spreadsheet file.
	Displays online help for the current report along with a complete outline of the I-SPY online help.
	Closes the report window.

Table 6.1 I-SPY Report Icons

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](#)

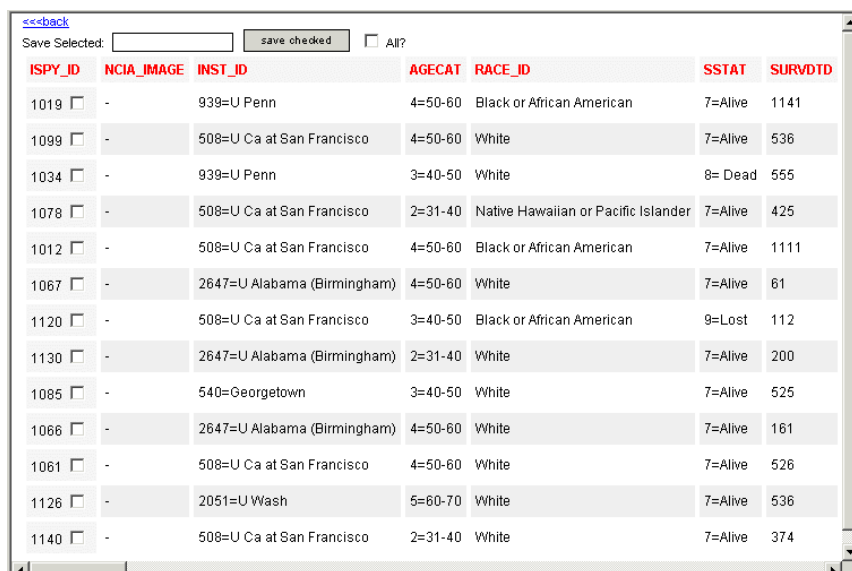
View Results (*Figure 6.1*) displays the query name and lists the output generated for the query.

Class Comparison (CC) (elapsed time: 861ms)	completed ✓
PCA (PCA) (elapsed time: 4756ms)	completed ✓
Search (Clinical) (elapsed time: 0ms)	completed ✓
C) Class Comparison (HC) Hierarchical Clustering (PCA) Principal Component Analysis	

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 a clinical report, you can create a [PatientDID list](#) to further filter queries. If there is an icon in the NCIA Image column of the report, you can also access images associated with a patient.



ISPY_ID	NCIA_IMAGE	INST_ID	AGECAT	RACE_ID	SSTAT	SURVDTD
1019	-	939=U Penn	4=50-60	Black or African American	7=Alive	1141
1099	-	508=U Ca at San Francisco	4=50-60	White	7=Alive	536
1034	-	939=U Penn	3=40-50	White	8=Dead	555
1078	-	508=U Ca at San Francisco	2=31-40	Native Hawaiian or Pacific Islander	7=Alive	425
1012	-	508=U Ca at San Francisco	4=50-60	Black or African American	7=Alive	1111
1067	-	2647=U Alabama (Birmingham)	4=50-60	White	7=Alive	61
1120	-	508=U Ca at San Francisco	3=40-50	Black or African American	9=Lost	112
1130	-	2647=U Alabama (Birmingham)	2=31-40	White	7=Alive	200
1085	-	540=Georgetown	3=40-50	White	7=Alive	525
1066	-	2647=U Alabama (Birmingham)	4=50-60	White	7=Alive	161
1061	-	508=U Ca at San Francisco	4=50-60	White	7=Alive	526
1126	-	2051=U Wash	5=60-70	White	7=Alive	536
1140	-	508=U Ca at San Francisco	2=31-40	White	7=Alive	374

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](#).

- 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*).



ISPY_ID	NCIA_IMAGE	INST_ID	AGECAT	RACE_ID	SSTAT	SURVDTD
1104	-	2527=U Texas SW Med Ctr	4=50-60	Black or African American	7=Alive	516
1128	-	508=U Ca at San Francisco	2=31-40	White	7=Alive	416
1053	-	372=U North Carolina at Chapel Hill	5=60-70	White	7=Alive	723
1009	-	508=U Ca at San Francisco	3=40-50	Asian	7=Alive	1021

Figure 6.3 Checking the I-SPY ID column

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



[<<<back](#)
 Save Selected: ☒ All?

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*).

<<back
Save Selected: ☒ All?
ISPY_ID **NCIA_IMAGE** **INST_ID**

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 side bar at the bottom of the PatientDID names (*Figure 6.6*). Mouse over the name and the data items appear.

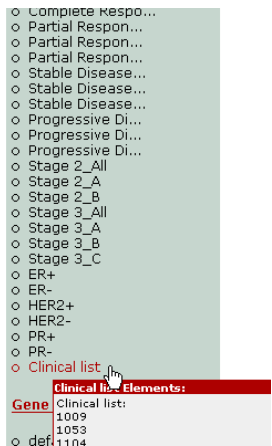



Figure 6.6 Saved PatientDID list

Accessing an NCIA Image from a Clinical Report

To access an NCIA image from a clinical report, follow these steps:

Click the NCIA icon  in the **NCIA Image** column (*Figure 6.7*).

Save Selected: ☐ All?



ISPY_ID	LabTrak ID	NCIA_IMAGE	INST_ID	AGECAT	RACE_ID	SSTAT	SURV
1035	214064		530=Mem Sloan-Kettering Cancer Ctr	3=40-50	White	9=Lost	177
1069	227882	-	508=U Ca at San Francisco	4=50-60	White	7=Alive	588
1110	227433	-	530=Mem Sloan-Kettering Cancer Ctr	4=50-60	White	7=Alive	371

Figure 6.7 Clinical Report with NCIA Image icon

The [NCIA](#) web site appears. As a first time user, register with NCIA to obtain a username and password. Then you can log in and have access to these images.

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](#) to further filter queries. You can also sort the patients using the red triangles in the column name.

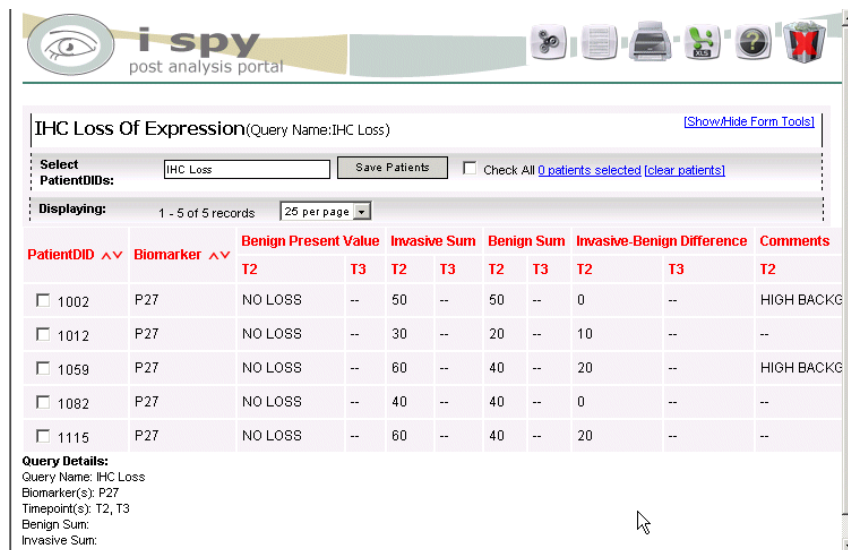


PatientDID	Biomarker	% Positive		Intensity		Localization		Result Code		Sum
		T2	T3	T2	T3	T2	T3	T2	T3	
1001	HER2	1-9	--	NEGATIVE	--	NONE	--	--	--	NO
1002	HER2	51-75	--	BORDERLINE	--	MEMBRANE AND CYTOPLASM	--	--	--	NO
1004	HER2	1-9	--	NEGATIVE	--	NONE	--	--	--	NO
1005	HER2	1-9	--	NEGATIVE	--	NONE	--	--	--	NO
1008	HER2	10-25	--	WEAK	--	MEMBRANE	--	--	--	NO
1009	HER2	1-9	--	NEGATIVE	--	NONE	--	--	--	NO
1010	HER2	1-9	--	NEGATIVE	--	NONE	--	--	--	NO
1012	HER2	1-9	--	NEGATIVE	--	NONE	--	--	--	NO

Figure 6.8 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](#) to further filter queries. You can also sort the patients using the red triangles in the column name.



PatientDID	Biomarker	Benign Present Value		Invasive Sum		Benign Sum		Invasive-Benign Difference		Comments
		T2	T3	T2	T3	T2	T3	T2	T3	
1002	P27	NO LOSS	--	50	--	50	--	0	--	HIGH BACKG
1012	P27	NO LOSS	--	30	--	20	--	10	--	--
1059	P27	NO LOSS	--	60	--	40	--	20	--	HIGH BACKG
1082	P27	NO LOSS	--	40	--	40	--	0	--	--
1115	P27	NO LOSS	--	60	--	40	--	20	--	--

Query Details:
 Query Name: IHC Loss
 Biomarker(s): P27
 Timepoint(s): T2, T3
 Benign Sum:
 Invasive Sum:
 Loss Result: NO LOSS

Figure 6.9 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 6.10](#)) displays the query name and lists the output generated for the query.

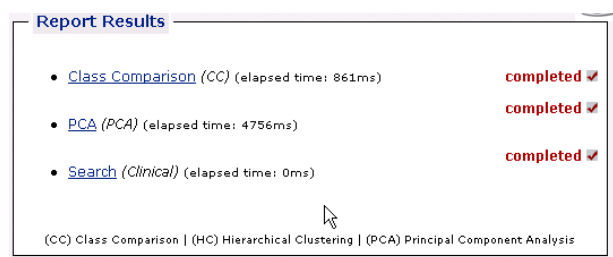


Figure 6.10 HOA Results

Class Comparison Report

The *Class Comparison* report ([Figure 6.11](#)) 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 6.11](#)), 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.

Reporter	Group Avg	P-Value	Fold Change	Gene Symbol
AGI_HUM1_OLIG0_A_23_P38830	-0.0806 / 1.6300	5.68e-6	-3.2730	ZNF552
AGI_HUM1_OLIG0_A_24_P408047	0.0679 / -1.0488	1.35e-3	2.1684	PLEKHA4
AGI_HUM1_OLIG0_A_24_P266048	-0.0444 / 1.1086	1.48e-3	-2.2238	FLJ20152
AGI_HUM1_OLIG0_A_24_P350759	-0.1635 / 0.9900	2.14e-3	-2.2246	SLC1A2
AGI_HUM1_OLIG0_A_23_P47614	0.1494 / -1.0800	2.77e-3	2.3448	PHLDA2

Figure 6.11 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](#). To create a Gene list, follow these steps (*Figure 6.12*):

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.
3. 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.

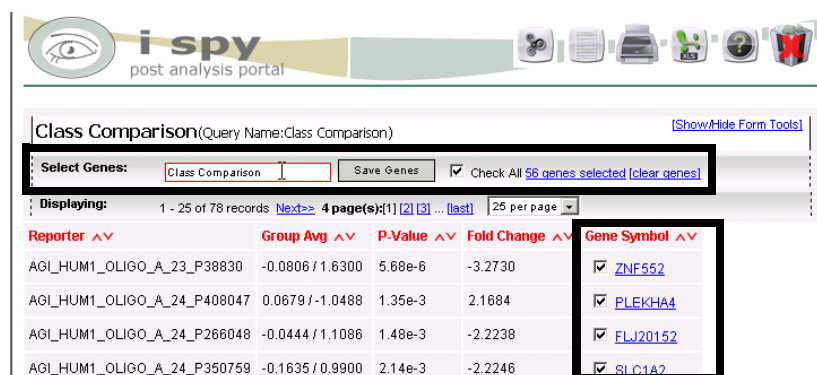


Figure 6.12 Selecting Genes instructions

4. Click the **Save Genes** button.

The results are saved.

5. Click the **OK** button.

The new Gene list appears in red in the side bar 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 6.13*).



Figure 6.13 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 6.14*).

Gene	Reporter
EGFR	1565483_at
EGFR	1565484_x_at

Figure 6.14 The Gene column

The Cancer Genome Anatomy Project (CGAP) browser opens.

Principal Component Analysis Plot

The *Principal Component Analysis plot* (*Figure 6.15*) is a two-dimensional graph which plots the various principal components from the analyses.

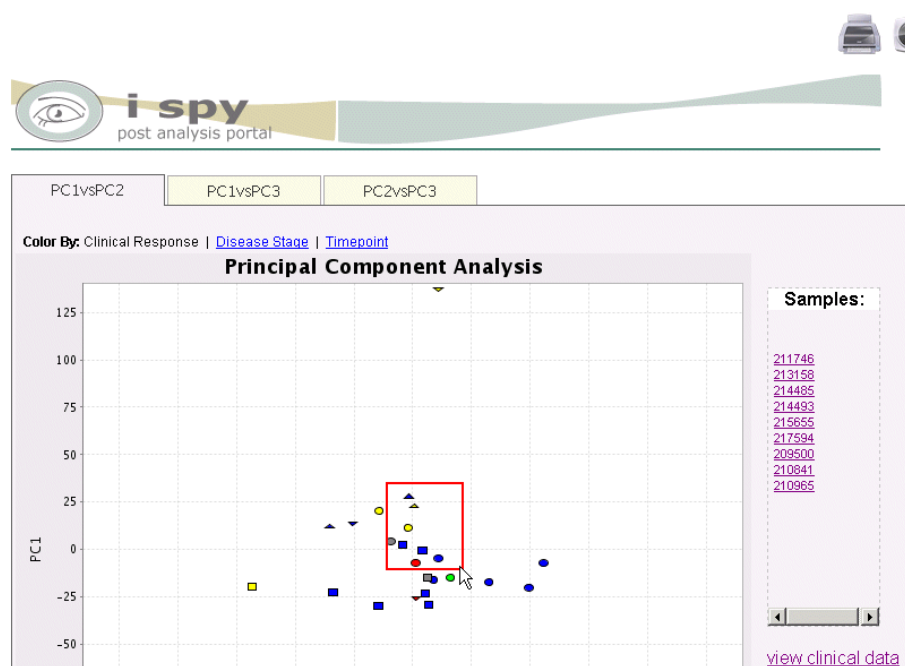


Figure 6.15 Principal Component Analysis Plot

Table 6.2 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: <ul style="list-style-type: none"> • PC1vsPC2 • PC1vsPC3 • PC2vsPC3
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 in 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 Selecting Samples of Interest in a Plot).

Table 6.2 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 6.16*).

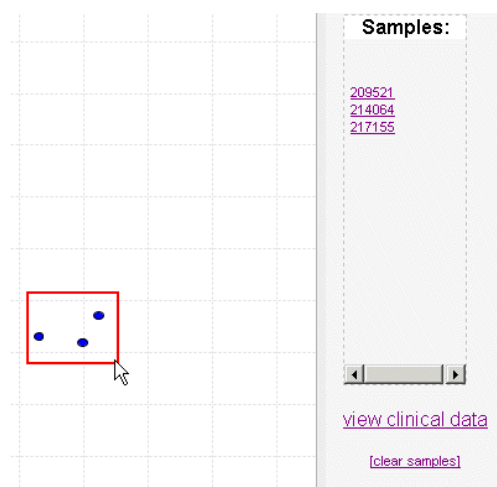


Figure 6.16 Sampling from a clinical plot

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 6.17*).

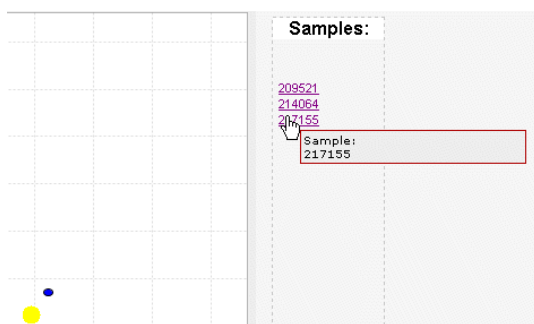


Figure 6.17 Lasso the points

3. To generate clinical data for the selected samples and save the samples, click the **view clinical data** link. See [Clinical Reports](#).

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

Hierarchical Clustering Report

The *Hierarchical Clustering report* (*Figure 6.18*) 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.

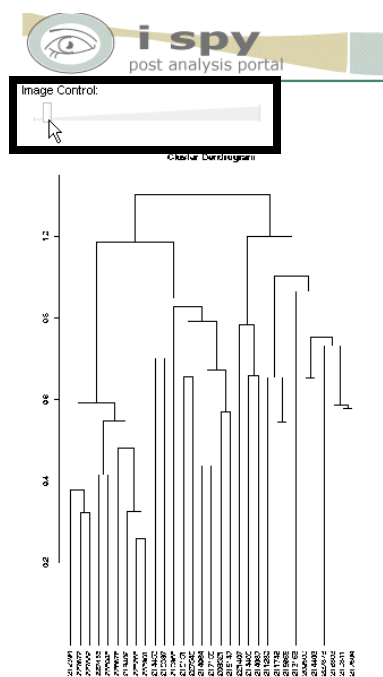


Figure 6.18 Hierarchical Clustering Dendrogram

A clinical report appears beneath the dendrogram where you can create a [PatientDID list](#) or display an image associated with a patient. For more information, see [Clinical Reports](#).

Save Selected: ☐ All?


ISPY_ID	LabTrak ID	NCIA_IMAGE	INST_ID	AGECAT	RACE_ID	SSTAT	SURV
1035 <input type="checkbox"/>	214064		530=Mem Sloan-Kettering Cancer Ctr	3=40-50	White	9=Lost	177
1069 <input type="checkbox"/>	227882	-	508=U Ca at San Francisco	4=50-60	White	7=Alive	588
1110 <input type="checkbox"/>	227433	-	530=Mem Sloan-Kettering Cancer Ctr	4=50-60	White	7=Alive	371

Figure 6.19 Hierarchical Clustering Clinical Report

Correlation Scatter Plot

The *Correlation Scatter plot* (Figure 6.20) 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.

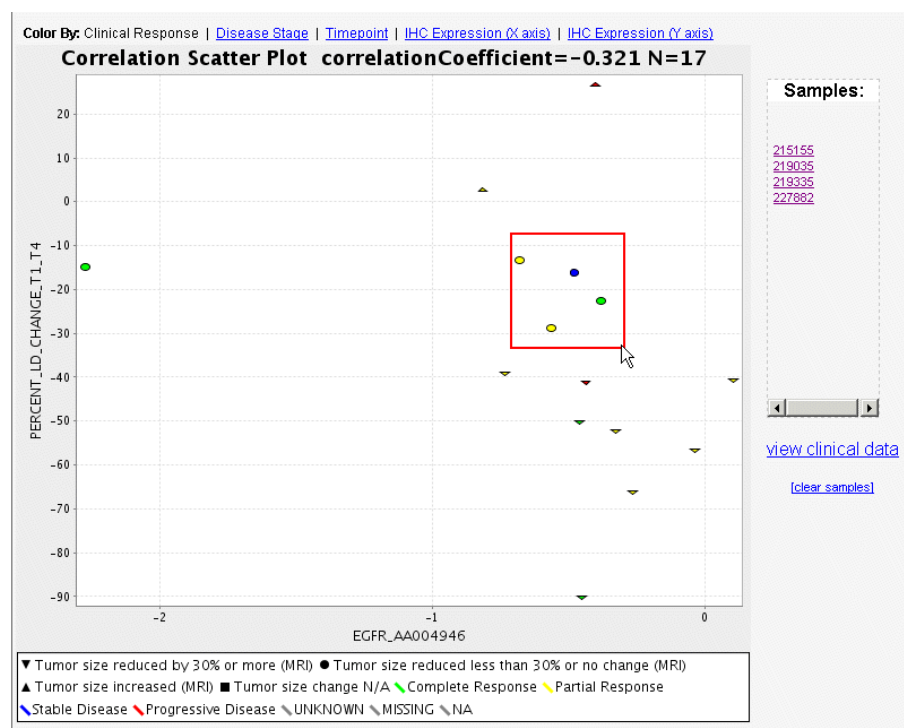


Figure 6.20 Correlation Scatter Plot

Table 6.3 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.
N=	The number of samples in the plot.
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 Selecting Samples of Interest in a Plot).

Table 6.3 Areas of the Correlation Scatter Plot

Categorical Plot Analysis

The *Categorical Plot analysis* (Figure 6.21) 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.

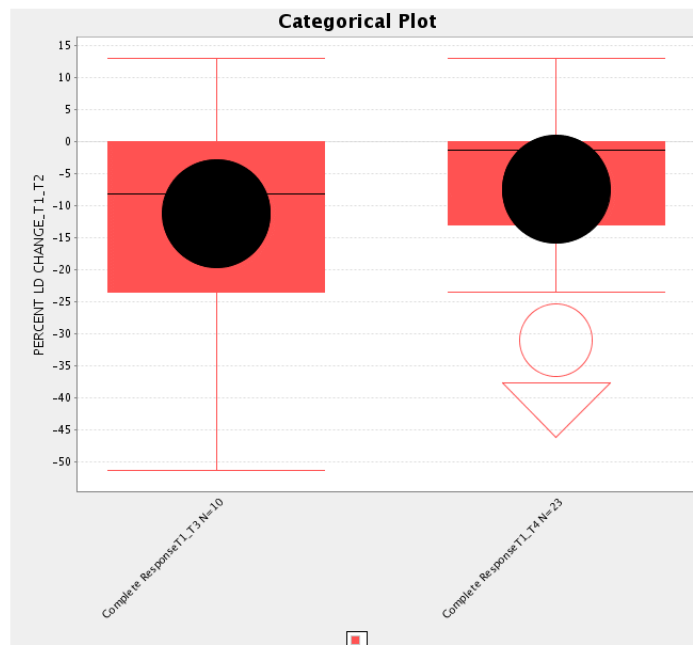


Figure 6.21 Categorical Plot

CHAPTER 7

MANAGING LISTS

This chapter describes how to manage patient and gene lists.

Topics in this chapter include:

- *Managing Lists Overview* on page 47
- *Viewing the Data Items in a List* on page 48
- *Removing Data Items to Create a New List* on page 49
- *Deleting an Entire List* on page 50
- *Adding a New “Custom” List* on page 51
- *Combining Existing Lists to Create a New List* on page 49

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 side bar displays all PatientDID lists. Mouse over a list name in the side bar to display the list's data items. Note that user-defined lists appear in red type in the side bar.

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 side bar displays the names of the Gene lists. Mouse over a list name in the side bar to display the list's data items. Note that user-defined lists appear in red type in the side bar.

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**).

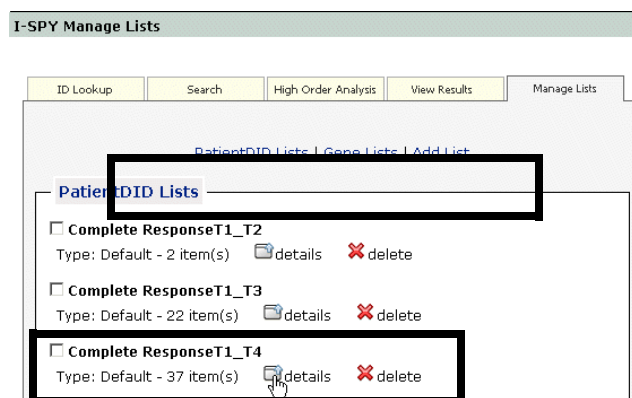


Figure 7.1 List types and Details

2. 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 side bar 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.

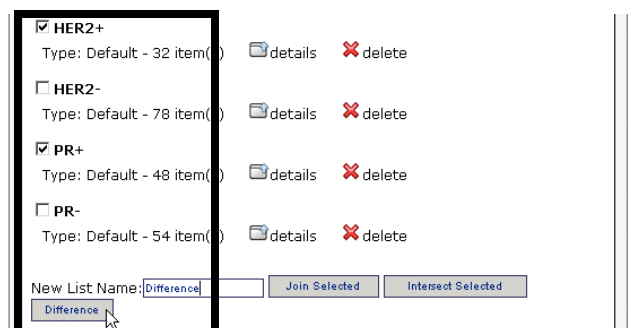


Figure 7.3 Combining existing lists

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*).

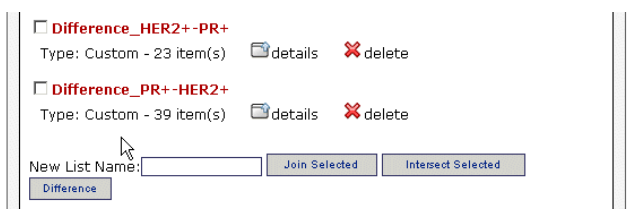


Figure 7.4 New Difference lists

The new list(s) appear on the Manage Lists page and in the side bar 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.

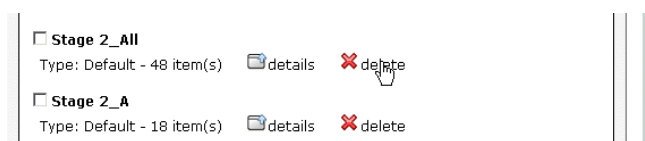


Figure 7.5 Deleting an entire list

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.

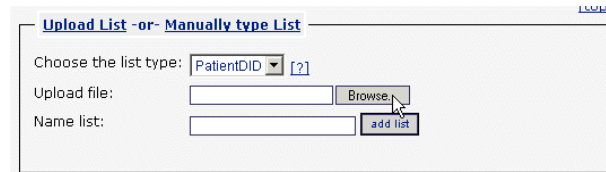


Figure 7.6 Uploading a list

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 side bar in red.
3. To create and add a list manually, follow these steps:
 1. Click **Manually Type List** at the top of the box.

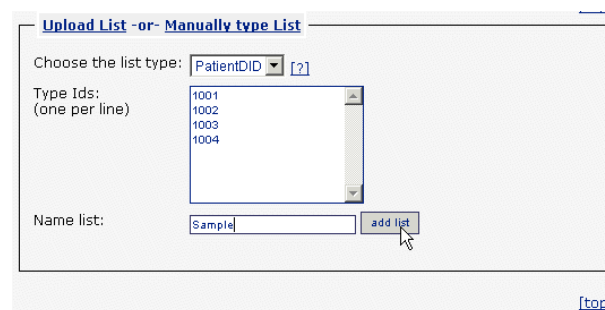


Figure 7.7 Manually typing a list

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 side bar in red.
5. To open and view the newly created list or save it to your computer, click on the list name in the side bar. 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 side bar. Click **Open** or **Save**.

APPENDIX A DATA DICTIONARY

This appendix lists the contents of the data dictionary.

Topics in this appendix include:

- *I-SPY Data Dictionary* on page 54
- *Patient Demographic Data Dictionary* on page 54
- *Chemotherapy Summary Data Dictionary* on page 56
- *On-Study Data Dictionary* on page 57
- *Post-Surgery Summary Data Dictionary* on page 58
- *Follow-Up Data Dictionary* on page 60
- *Response Evaluation Data Dictionary* on page 60

I-SPY Data Dictionary

Table A.1 lists the variables, description, and format for the data dictionary:

Variable Name	Variable Description	Format
ISPY_ID (or De_ID)	I-SPY identifier uniquely corresponds 1 to 1 to the CALGB patient identifier	Integer 1001-1239
DataExtractDt	Date data was downloaded from the CALGB database, for this transfer it is April 16, 2007	Date 10 format mm/dd/yyyy
Inst_ID	Registering Institution 372=U North Carolina at Chapel Hill 443=Chicago 508=U Ca at San Francisco 530=Mem Sloan-Kettering Cancer Ctr 540=Georgetown 2527=U Texas SW Med Ctr 939=U Penn 2051=UWash 2527=UTexas 2647=U Alabama (Birmingham) 2790=ECOG	Num

Table A.1 I-SPY variables

Patient Demographic Data Dictionary

Table A.2 lists Patient Demographic variables, description, and format for the data dictionary:

Variable Name	Variable Description	Format
ISPY_ID (or De_ID)	I-SPY identifier uniquely corresponds 1 to 1 to the CALGB patient identifier	Integer 1001-1239
DataExtractDt	Date data was downloaded from the CALGB database, for this transfer it is April 16, 2007	Date 10 format mm/dd/yyyy

Table A.2 Patient demographics variables

Variable Name	Variable Description	Format
Inst_ID	Registering Institution 372=U North Carolina at Chapel Hill 443=Chicago 508=U Ca at San Francisco 530=Mem Sloan-Kettering Cancer Ctr 540=Georgetown 2527=U Texas SW Med Ctr 939=U Penn 2051=UWash 2527=UTexas 2647=U Alabama (Birmingham) 2790=ECOG	Num
AgeCat	Patient Age Category 1= 18-30 2= >30-40 3= >40-50 4= >50-60 5= >60-70 6= >70-80 7= >80-<89 >=89 not available	Integer
Age	Patient Age	Integer
Race_id	Patient Race (1 case is 136 for white, black and American Indian or Alaska Native)	Character
Sstat	Survival Status (7=Alive, 8= Dead, 9=Lost)	Integer
SurvDtD	Survival date (time from study entry to death or last follow-up)	Integer

Table A.2 Patient demographics variables

Chemotherapy Summary Data Dictionary

Table A.3 lists Chemotherapy Summary data variables, description, and format for the data dictionary:

Variable Name	Variable Description	Format
Chemo A=Adria C=Cytosan E=Epirubicin F=5-Fu Td=Docetaxel Tp=Paclitaxel H=Herceptin TpTd =Crossover from Tp to Td or Td to Tp	Neo-Adjuvant Chemotherapy Regimen 1=AC 2=EC 3=FEC 4=FAC 5=A 6=AC → Td 7=AC → Tp 8=A → Td 9=A → Tp 12=AC → TdH 13=AC → TpH 14=FEC → Tp 15=EC → Tp 16=AC → TpTd 17= A → Tp → C 19=AC → Td → Xeloda 20=EC → Tp → Carboplatin 21=FEC → Tp → Abraxane 22=AC → Td → Navelbine → Xeloda 23=AC → Tp → Vinorelbine → Tarceva 24= C → Tp	Integer
ChemoCat	Chemotherapy Group Category (Condensed from prior variable "Chemo") 1=Anthracycline Only Regimen 2=Anthracycline Plus Taxane 4=Anthracycline + Taxane + Other	Integer
DoseDenseAnthra	Dose Dense Anthracycline Therapy? 0= Standard Therapy, q3 wks 1= Dose Dense Therapy, q2 wks	Integer
DoseDenseTaxane	Dose Dense Taxane Therapy? 0= Standard Therapy, q3 wks 1= Dose Dense Therapy, q2 wks	Integer
Tam	Tamoxifen received 1=Yes	Integer

Table A.3 Chemotherapy summary data variables

Variable Name	Variable Description	Format
Herceptin	Herceptin received 1=Yes	Integer

Table A.3 Chemotherapy summary data variables

On-Study Data Dictionary

Table A.4 lists On-Study data variables, description, and format for the data dictionary:

Variable Name	Variable Description	Format
MenoStatus	Menopausal Status 1=Pre (<6 mo since LMP AND no prior bilateral ovariectomy AND not on estrogen replacement) 2= Post (prior bilateral ovariectomy OR > 12 mo since LMP with no prior hysterectomy) 5=Indeterminate, not 1 or 2 above	Integer
SentinelNodeSample	Was Sentinel node sampling performed pre-treatment? 0=no 1=yes	Integer
SentinelNodeResult	Sentinel Node Biopsy Results pre-treatment 0=Negative 1=Positive	Integer
HistTypeInv	Histologic Type of Invasive Tumor (On-Study) 1=Necrosis 2=Ductal Carcinoma 3=Lobular 4=Mixed Ductal/Lobular carcinoma 5=Other 6=No invasive tumor present	Integer
HistologicGradeOS	Combined Histologic Grade - On-study (According to SBR/Elston Classification) 1=Grade I (low) 2= Grade II (intermediate) 3= Grade III (high) 4= Indeterminate	Integer
ER_TS	Estrogen Receptor Status – Total Score Total Score = ER_PS+ ER_IS Considered Allred Score; 3 is positive	Integer

Table A.4 On-Study summary data variables

Variable Name	Variable Description	Format
PgR_TS	Progesterone Receptor Status – Total Score Total Score = PgR_PgS+ PgR_IS Considered Allred Score; 3 is positive	Integer
The following Her2 measures were performed in the community as opposed to centrally If FISH is 3+.. consider positive; if IHC is 2+ look at FISH result		
Her2CommunityPos (Based on Algorithm of the Her2 community measured variables)	Her2 Summary as measured in the Community 0=negative 1=positive* 2= indeterminate** 3=not done *Any 3+ or FISH ratio 2 ** 2+	Integer
Her2CommunityMethod	Her2 Summary method as measured in the Community 1= IHC 2=FISH 3= Other 4=Unknown	Integer

Table A.4 On-Study summary data variables

Post-Surgery Summary Data Dictionary

Table A.5 lists Post-Surgery Summary data variables, description, and format for the data dictionary:

Variable Name	Variable Description	Format
SurgeryLumpectomy	Surgery Procedure Performed was Partial mastectomy/ lumpectomy/ excisional biopsy 0=no 1=yes	Integer
SurgeryMastectomy	Surgery Procedure Performed was Mastectomy, NOS 0=no 1=yes	Integer
InitLump_FupMast	Initial Lumpectomy Surgery followed by Mastectomy Surgery at a later date 1=Yes	Integer

Table A.5 Post-surgery summary data variables

Variable Name	Variable Description	Format
Surgery	Did patient have extensive Primary Surgery immediately following chemotherapy? 1=Yes, Lumpectomy/Mastectomy or Both 0=Delayed surgery (>8 weeks); following RT 2= Delayed surgery (>8 weeks); other than primary RT 3=No, Disease Progression 4=No, Adverse Event 5=No, Patient Refusal 6=No, Other Reason, Not specified	Integer
DCISonly	DCIS only thing left following surgery 1=Yes	Integer
PTumor1Szcm_Micro	Primary Tumor Pathological Tumor Size – Microscopic, measured in cm	Number with decimal
HistologicTypePS	Histologic Type of Primary Tumor (Post-Surgery) 1=Necrosis 2=Ductal Carcinoma 3=Lobular 4=Mixed Ductal/Lobular carcinoma 5=Other 6=No residual invasive breast cancer	Integer
HistologicGradePS	Combined Histologic Grade -Post-Surgery (According to SBR/Elston Classification) 1=Grade I (low) 2= Grade II (intermediate) 3= Grade III (high) 4= Indeterminate 5=No residual invasive breast cancer	Integer
NumPosNodes	Total Number positive Axillary + Sentinel (post) Nodes, post-chemotherapy	Number
NodesExamined	Total Number of Axillary + Sentinel (post) nodes Examined, post-chemotherapy	Number

Table A.5 Post-surgery summary data variables

Variable Name	Variable Description	Format
PathologyStage	Pathology Assessment Staging 1= Stage 0 (DCIS only) 2= Stage I 3= Stage IIA 4=Stage IIB 5= Stage IIIA 6= Stage IIIB 7=Stage IIIC 8= Stage IV	Integer
ReasonNoSurg	Principal Reason Breast Conserving Surgery Not Performed 1=Multicentric Disease 2=Inflammatory Disease 3=Diffuse microcalcifications 4=Patient Choice/ Family history 5=Institutional Norm 6=Specific anatomy of primary 7=Other (Note this data comes from both the C-911 and C-931 forms)	

Table A.5 Post-surgery summary data variables

Follow-Up Data Dictionary

Table A.6 lists Follow-Up Summary data variables, description, and format for the data dictionary:

Variable Name	Variable Description	Format
RtTherapy	Has patient received adjuvant radiation therapy (prior to treatment failure or second primary cancer)? 1=No , 2= Yes	Integer
RtBreast	Radiation to Breast 0= No, 1 = Yes	Integer
RtBoost	Radiation to Boost 0= No, 1 = Yes	Integer
RtAxilla	Radiation to Axilla 0= No, 1= Yes	Integer
RtSNode	Radiation to Supraclavicular node 0= No, 1= Yes	Integer
RtIMamNode	Radiation to Internal Mammary node 0=No, 1 = Yes	Integer

Table A.6 Follow-up summary data variables

Variable Name	Variable Description	Format
RTChestW	Radiation to Chest Wall 0=No, 1= Yes	Integer
RtOther	Radiation to Other Site 0= No, 1= Yes	Integer

Table A.6 Follow-up summary data variables

Response Evaluation Data Dictionary

Table A.7 lists Response Evaluation data variables, description, and format for the data dictionary:

Response evaluation form coincides with MRI schedule.

- T2 = Early Treatment Day1, cycle 2
- T3= Inter-Regimen
- T4= Prior to Surgery for Response evaluation form and sample Post-surgery

Variable Name	Variable Description	Format
TSizeClinical	Size of Primary Tumor (cm) – Clinical Assessment at Baseline	Number with decimal
NSizeClinical	Size of Largest Palpable Node (cm) – Clinical Assessment at Baseline	Number with decimal
StageTe	Disease Stage T (metastasis) Baseline	Char
StageNe	Disease Stage N (metastasis) Baseline	Char
StageMe	Disease Stage M (metastasis) Baseline	Char
ClinicalStage (Based on algorithm)	Clinical Staging at Baseline 1= Stage 0 2= Stage I 3= Stage IIA 4=Stage IIB 5= Stage IIIA 6= Stage IIIB 7=Stage IIIC 8= Stage IV 9=Inflammatory	Integer
ClinRespT1_T2 (If missing from form used an algorithm to determine) RECIST criteria	Clinical Response Baseline to Early Treatment 1= CR 2=PR 3=Stable Disease 4=Progressive Disease	Integer

Table A.7 Response Evaluation summary data variables

Variable Name	Variable Description	Format
ClinRespT1_T3 (If missing from form used an algorithm to determine) RECIST criteria	Clinical Response Baseline to Inter-Regimen 1= CR 2=PR 3=Stable Disease 4=Progressive Disease 5=Inter-Regimen Chemo not given	Integer
ClinRespT1_T4 (If missing from form used an algorithm to determine) RECIST criteria	Clinical Response Baseline to Pre-Surgery 1= CR 2=PR 3=Stable Disease 4=Progressive Disease 6=No Surgery Performed	Integer

Table A.7 Response Evaluation summary data variables

GLOSSARY

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

Term	Definition
CGAP	Cancer Genome Anatomy Project
Class Comparison	Differential gene expression across the tumor types will be evaluated by calculating the typical t -statistic for each reporter. Both parametric and non-parametric p -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. These are shown in the Manage Lists function.
Group	A pre-defined or user-defined PatientDID list comprising patient identifiers with certain characteristics.
Hierarchical Clustering	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.

Table B.1 Glossary of I-SPY terms

<i>Term</i>	<i>Definition</i>
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.
NCIA	National Cancer Imaging Archive
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. These are shown in the Manage Lists function.
Principal Component Analysis	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, for example defining parts, while simultaneously filtering out noise.

Table B.1 Glossary of I-SPY terms

INDEX

Symbols

39

A

Across Timepoints, definition 22

Adding

custom list (uploading or typing) 51

new list (combining) 49

new list (removing data items) 49

user account 4

Analyses

Categorical Plot 30

Class Comparison 22

Correlation Scatter Plot 28

Hierarchical Clustering 26

Principal Component 24

Audience

of the User's Guide v

Average linkage, definition 27

B

Baseline, assigning 23

Black dot in box

on Categorical Plot 44

Blue panel

description 6

red name user created 36

red PatientDID list (Clinical report) 36

Box and Whisker plot

Categorical Plot 30

C

Categorical Plot

Analysis, description 30

Analysis, report, description 44

Analysis Form,description 30

black dot in box defined 44

circles defined 44

High Order Analysis 30

line in box defined 44

triangles defined 44

uses for 44

uses for Box and Whisker plot 30

using Managing Lists 30

Circles

on Categorical Plot 44

Class Comparison

Analysis, definition 22

Analysis Form,description 22

example (View Results) 38

High Order Analysis 22

report description (View Results) 38

Clinical plot

selecting samples of interest 41

Clinical Query

definition 13

filling in form 14

Form page, description 14

Clinical report

creating PatientDID list 35

description 35

saving genes 39

saving samples 36

selecting samples 35

Closing

a report 34

Combining lists 49

Complete linkage, definition 27

Constrain by GeneList, definition 25, 26

Constrain reporters by variance (Gene Vector)

percentile, definition 25, 26

Correlation, definition 29

Correlation Coefficient, definition 44

Correlation Scatter Plot

Analysis, definition 27

Analysis Form,description 28

examples of 27

High Order Analysis 28

report, description 43

Creating

custom list (uploading or typing) 51

new list (combining) 49

- new list (removing data items) 49
- user account 4

D

- Data items
 - removing from a list 49
 - viewing in a list 48
- defaultGene1, description 26
- Deleting
 - list 50
- Difference option, description 50
- Distance Matrix, definition 27
- Downloading
 - ID Lookup data 12

E

- E-mailing
 - NCICB Application Support 4
- Euclidean distance, definition 27
- Excel file
 - downloading ID Lookup data 12

F

- Fixed Timepoint, definition 22
- Fold Change, definition 23

G

- Gene list
 - creating with Manage Lists 49
 - saving on Class Comparison report 39
 - selecting and saving 39
- Group
 - selecting 14

H

- Help
 - obtaining 7
- Help link 7
- Hierarchical Clustering
 - Analysis, definition 26
 - Analysis Form, description 26
 - dendogram description 43
 - High Order Analysis 26
 - NCIA icon 36
 - page description (View Results) 42
 - report, description 42
- High Order Analysis function
 - Categorical Plot 30
 - Class Comparison 22
 - Correlation Scatter Plot 28
 - Hierarchical Clustering 26
 - overview 21

- Principal Component 24
- Horizontal line in box
 - on Categorical Plot 44

I

- ID Lookup function
 - creating PatientDID list 12
 - displaying patient sample information 11
 - downloading data 12
 - overview 9
 - steps 10
- IHC Expression X/Y axis, Color By, definition 44
- IHC Level of Expression
 - defined 16
 - filling in Query form 16
 - report, description 37
 - report, example 37
- IHC Loss of Expression
 - defined 18
 - filling in Query form 18
 - Form page, description 18
 - report, description 37
 - report, example 37
- Image Control bar 42
- Intersect option, description 50
- I-SPY
 - creating user account 4
 - description 1
 - functions 2
 - getting help 7
 - launching 3
 - logging in 5
 - logging out 8
 - menu 7
 - workspace description 6
 - workspace screen 6
- ISPY ID
 - definition 11

J

- Join option, description 50

L

- LabTrak ID
 - definition 11
- Launching
 - I-SPY 3
- Legal Rules of the Road page 5
- Line in box
 - on Categorical Plot 44
- Linkage Method, definition 27
- Lists
 - adding a custom list 51

- combining lists [49](#)
- creating a new list [49](#)
- deleting [50](#)
- removing data items [49](#)
- viewing data items [48](#)

Logging in

- I-SPY [5](#)

Logging out

- I-SPY [8](#)

Logout link [8](#)

M

Managing Lists function

- creating PatientDID list [35](#)
- for Categorical Plot analysis [30](#)
- overview [47](#)

Manually Type List options, description [51](#)

Microsoft Excel

- downloading ID Lookup data [12](#)

Morphology, definition [15](#)

N

NCIA icon [36](#)

O

Overview

- High Order Analysis function [21](#)
- ID Lookup function [9](#)
- managing lists [47](#)
- Search function [13](#)
- View Results function [33](#)

P

PatientDID list

- creating with ID Lookup [12](#)
- creating with Manage Lists [49](#)
- description [48](#)
- in Clinical report [35](#)
- saving on Clinical report [36](#)
- user list blue panel [36](#)

Patient ID

- displaying sample information [11](#)
- searching for [10](#)

Pearson correlation, definition [27](#), [29](#)

Percent LD, definition [15](#)

PERCENT LD_CHANGE, definition [29](#), [31](#)

Principal Component

- Analysis, defined [24](#)
- Analysis Form,description [24](#)
- Analysis report, description [40](#)
- High Order Analysis [24](#)
- report example [40](#)

Printing

- a report [34](#)

Purpose

- of the User's Guide [v](#)

p-value, definition [23](#)

R

Red list name

- user created list [36](#)

Removing

- data items in a list [49](#)

Reports

- Categorical Plot [44](#)
- Class Comparison [38](#)
- Correlation Scatter Plot report example [43](#)
- Hierarchical Clustering [42](#)
- Principal Component Analysis [40](#)
- using icons [34](#)
- viewing [33](#)

S

Sample

- selecting on Clinical report [35](#)
- selecting on plot [41](#)

Sample ID

- displaying patient information [11](#)
- searching for [10](#)

Samples area

- selecting samples [41](#)

Samples of interest

- [41](#)

Saving

- a report [34](#)
- genes on Class Comparison [39](#)
- samples on Clinical report [36](#)

Search function

- Clinical Query [14](#)
- IHC Level of Expression Query [16](#)
- IHC Loss of Expression Query [18](#)
- overview [13](#)

Searching

- patient IDs and samples [10](#)

Select Genes toolbar [39](#)

Select Group

- assigning [14](#)

Side bar

- manage list information [48](#)
- red Gene list (Class Comparison) [39](#)
- red PatientDID list (Search) [12](#)

Single linkage, definition [27](#)

Sorting

- report columns [39](#)

Spearman, definition [29](#)

Spreadsheet

saving report to [34](#)

Statistical Method, definition [23](#)

Support link [7](#)

T

Timepoint, definition [11](#)

Triangles

on Categorical Plot [44](#)

top of columns to sort [39](#)

T-test Two Sample Test

definition [23](#)

output [38](#)

Typing

a list (Manage List function) [51](#)

U

Uploading

list (Manage List function) [51](#)

User's Guide

audience [v](#)

organization of [v](#)

purpose [v](#)

text conventions [vi](#)

User account

creating [4](#)

User guide link [7](#)

V

Viewing

data items in a list [48](#)

View Results function

overview [33](#)

W

Wilcoxon Test Man-Whitney Test

definition [23](#)

output [38](#)

Y

Yellow circle

meaning on plot [42](#)

