

## ICRP: adding value to cancer research

ICRP's mission is to add value to cancer research efforts internationally by fostering collaboration and strategic co-ordination between cancer research organizations.

Lynne Davies ICRP Operations Manager

Tel. +44 788 959 9948

lynne.davies@cancer.org.uk

22<sup>nd</sup> November, 2011



#### Milestones

CSO classification developed

7 organizations

• First generation website launched

• Agreement to share data between Partners

New website launched to Partners

Gives Partner access to full international dataset

• 51 member organizations

Public launch of ICRP site

Additional recruitment

2010

2000

2011

Next

HRCS developed from CSO



#### Directors & managers from the following organizations:

USA	Canada	Europe
American Cancer Society	Canadian Cancer Research Alliance (Consortium of 20 largest funders)	National Cancer Research Institute (UK Consortium of 19 largest funders)
American Institute for Cancer Research		Dutch Cancer Society
Avon Foundation		French National Cancer Institute
California Breast Cancer Research Program		
US Department of Defense (CDMRP)		
Susan G. Komen for the Cure		
National Cancer Institute		
National Pancreas Foundation		
Pancreatic Cancer Action Network		
Oncology Nursing Society Foundation		



#### How we work





## ICRP's web site provides

- A public site to allow users to search for research awards using defined criteria and is a valuable tool for researchers to identify potential collaborators worldwide
- Partner-only analytical tools on the new web site to allow organizations to conduct their own analyses of the international portfolio, giving our Partners an international perspective to help inform strategic planning
- Online networking tools and document exchange for our Partner organizations via a web site forum



Partners are able to access the restricted site from the public site

artner Login

HOME

**PARTNERS** 

PARTNER BENEFITS

DATABASE

CSO

CONTACT

SEARCH

Established in 2000, ICRP is a unique alliance of cancer organisations working together to enhance global collaboration and strategic coordination of research. We aim to improve access to information about cancer research being conducted and enable cancer organisations to maximise the impact of their independent efforts, for the benefit of researchers and cancer patients worldwide.

- ICRP includes organisations from Canada, France, the Netherlands, United Kingdom, and the United States.
- ICRP organisations share funding information in a common format (known as the Common Scientific Outline or CSO) to facilitate pooling data and evaluating data across organisations.
- The ICRP database contains information on more than 42,000 grants, totalling some \$XXXXX in cancer research from 48 organisations.
- Researchers can search the ICRP to avoid duplication and identify collaborators.

If you fund cancer research anywhere in the world, you should be a member of ICRP. Learn more about becoming a member.

#### 

#### **Dutch Cancer Society Joins ICRP**

The International Cancer Research Partners are pleased to announce addition of the Dutch Cancer Society (DCS) to the partnership.

#### 2010 ICRP Annual Meeting

The 2010 ICRP Annual Meeting will be held May 17–19, 2010 in Toronto, Canada.



International Cancer Research Partnership

Updated: May 14, 2010

Contact the Webmaster

Submit Feedback on this Page



Partners Site

SEARCH

HOME MylCRP

**DATA RESOURCES** 

Aberdeen, Scotland

Adelaide

**SEARCH & ANALYSIS** 

**FORUM** 

Brazil

Canada

LIBRARY

#### Search the ICRP Database

The ICRP database includes research awards from all member organisations, structured in an internationally recognised classification system known as the Common Scientific Outline (CSO). The database allows users to identify potential collaborators and avoid du Comprehensive used by members to find appropriate researchers to assist with peer reviewer of grant applications and journal article search functionality Enter search criteria below to search the ICRP database. All fields are option Search the database for awards containing: Users can search by all of these words: year, organization, this exact phrase: city, country, CSO, any of these words: cancer type, none of these words: keyword, project Search the Database type, PI Name.... Institution Receiving Award Institution Name: (Full or Partial) Principal Investigator: (Last Name) (First Name or Initial) City State/Territory Country All States All Cities All Countries ALBANY Alabama Argentina ANN ARBOR Alaska Australia ATLANTA Alberta Austria AURORA American Samoa Belgium

Arizona

Arkansas

Partners Site

HOME

**MyICRP** 

**DATA RESOURCES** 

**SEARCH & ANALYSIS** 

**FORUM** 

LIBRARY

SEARCH

#### **Database Search Results**

#### Search Criteria:

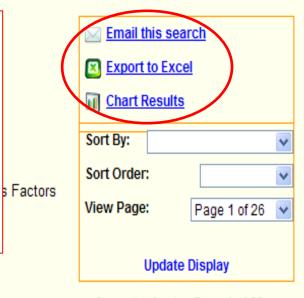
Funding Years: 2010, 2009

CSO Codes:

- · 2.1 Exogenous Factors in the Origin and Cause of Cand
- 2.2 Endogenous Factors in the Origin and Cause of Car
- 2.3 Interactions of Genes and/or Genetic Polymorphism
- · 2.4 Resources and Infrastructure Related to Etiology

Data analyses (including abstracts & \$ spend) can be exported to Excel or emailed

Organizationspecific caveats are included in the site



Your search returned 627 awards.

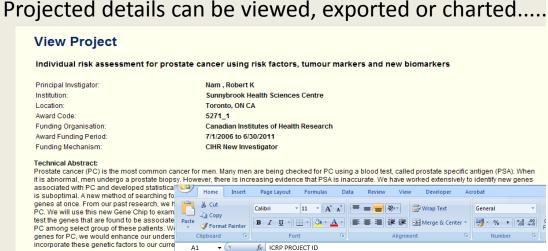
Currently Viewing Page 1 of 26

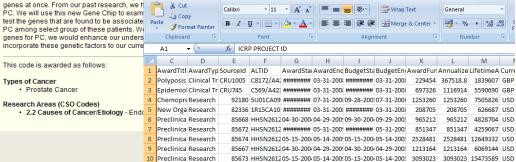
Title	Principal Investigator	Institution	City	State	Co.	Funding Org.	Award Code
Women's international study of long duration oestrogen after menopause (WISDOM)	Dunstan, D	St Bartholomew'S Hospital & Royal London School of Medicine & Dentistry	London, England		UK	DOH	DOH100
Elucidating And Modeling Irradiation Effects On Centrosomal And Chromosomal Stability Within Breast Cancer	Maxwell, Christopher	Lawrence Berkeley National Laboratory	Berkeley	CA	US	DOD, CDMRP	BC050612

Online, the results list can be sorted in ascending or descending order by any of the column headings. You can drill down to the detail on any award



#### Projected details can be viewed, exported or charted.....





11 Technical Research

12 Preclinica Research

13 Preclinica Research

14 Preclinica Research

15 Preclinica Research

17 Preclinica Research

19 Selective Research

20 Aspirin, U Research

21 Molecular Research

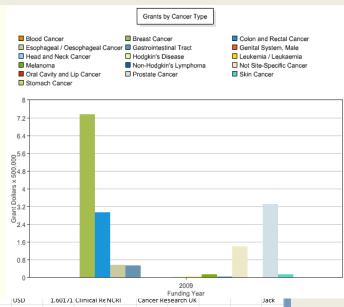
22 HMG Coa (Research

23 Mechanis Research

24 Methyl Se Research

25 Phytocher Research

18 2-Methox Research RSG-04-169-01



ι	J01 - Rese	NCI	National I	Division o	DCP	Tangrea, J	Arie S
F	R15 - Acad	NCI	National I	Division o	DCP	Perloff, M	Mark E
N	N01 - Rese	NCI	National I	Division o	DCP		David L
N	N01 - Rese	NCI	National I	Division o	DCP		John G
N	N01 - Rese	NCI	National I	Division o	DCP		Margie
N	N01 - Rese	NCI	National I	Division o	DCP		Martin
N	N01 - Rese	NCI	National I	Division o	DCP		Ming
N	NO2 - Resc	NCI	National I	Division o	DCP		Carolin
N	N01 - Rese	NCI	National I	Division o	DCP		Carol
N	N01 - Rese	NCI	National I	Division o	DCP		David
N	N01 - Rese	NCI	National I	Division o	DCP		Alex
N	N01 - Rese	NCI	National I	Division o	DCP		Keith A
N	NO2 - Resc	NCI	National I	Division o	DCP		Susan I
N	N01 - Rese	NCI	National I	Division o	DCP		Clintor
F	RSG - Rese	ACS	American	Cancer Soc	iety		Addanl
F	RO1 - Rese	NCI	National I	Division o	DCB	Yang, Shei	Bao-Tir
F	RO1 - Rese	NCI	National I	Division o	DCP	Richmond	John D
F	RO1 - Rese	NCI	National I	Division o	DCP	Maruvada	Clemei
F	RO1 - Rese	NCI	National I	Division o	DCP	Malone, V	Chinth
F	RO1 - Rese	NCI	National I	Division o	DCB	Snyderwir	Timoth
F	RO1 - Rese	NCI	National I	Division o	DCP	Perloff, M	Junxua
F	RO1 - Rese	NCI	National I	Division o	DCP	Perloff, M	Karen J

5590690 GBP

7505826 USD

626687 USD

4828704 USD

4961546 USD

208705

965212

0

356152

315043

842309

356152

85680 HHSN2612 09-30-200 09-29-200 09-30-200 09-29-200 2207214 2207214 11042117 USD

85677 HHSN2612 09-30-200 09-29-200 09-30-200 09-29-200 7415652 7415652 37098577 USD

85669 HHSN2612 04-30-200 04-29-200 04-30-200 04-29-200 1160472 1160472

85671 HHSN2612 06-30-200 06-29-200 06-30-200 06-29-200 1120048 1120048

85666 HHSN2612 04-30-200 04-29-200 04-30-200 04-29-200 1340679 1340679

######## 06-30-2001 ######## 06-30-2001 720000

85665 HHSN261204-30-200-04-29-200:04-30-200-04-29-200: 905285

96569 5R01CA09 05-13-200 03-31-200 ######## 03-31-200 842309

93813 5R01CA09 ######## 03-31-200: ######## 03-31-200: 286472

97792 5R01CA09 ######## 02-28-200\ ######## 02-28-200\ 260726

97961 5R01CA07 09-30-199 01-31-200 ######## 01-31-200 391069

85670 HHSN2612 ######## ######## ######## #########

107849 5R01CA09 ####### 02-28-200; ######## 02-28-200;

93423 5R01CA09 ######## 03-31-200( ######## 03-31-200)

93714 5R01CA09 ######## 03-31-200 ######## 03-31-200



## Capabilities of the ICRP database

The ICRP database contains information on more than 53,000 grants from 51 organizations and is expanding every month. We are using this unique resource to ask the questions

- What are the gaps in cancer research?
- What are the trends in cancer research funding?
- What types of projects have been funded (e.g. Clinical, Research, Training)?
- What research areas are being funded (CSO)?
- What types of cancer are being funded?
- What is the funding profile across different countries and organizations?
- How can we maximize our research efforts?
- How can we foster strategic collaborations between funding organizations?



#### Infrastructure

In-kind investment in the web site and database by NCI

1 part-time staff

Membership fees

- Provides a .org site but with federal government infrastructure
- NCI provides ongoing support, upgrades & maintenance through its contractor: NOVA Research
- •Any further modification can be planned and financed by the Partners
- Annual meeting
- Modifications/enhancements to the web site
- Operations Manager (50%fte)



## Challenges – web site/database

- (1) Setting up a common classification system
- (2) Acquiring and submitting full datasets
- (3) Agreeing a framework to share data

ICRP has put policies and procedures in place to safeguard organizations' data and to ensure that the data is used for the planned purpose. Key policies are that:

- Partners may use the data made available through the restricted view for internal purposes only.
- If Partners wish to publish any part of the data/reports on the restricted site, they must gain the approval of the Partnership first.
- Important caveats relating to Partners' data will be included on the site for reporting.
- New Partners cannot gain access to the international portfolio until they have contributed data and agreed to abide by the policies and procedures.



## Our activities: Databases...and beyond

ICRP offers a unique opportunity for cancer research funding agencies to share experiences and resources

- Portfolio analysis via the online web database and as collaborative groups
   e.g. Chemoprevention analysis, Sept. 2011
- Evaluations, led by Partner interests
  - e.g. Evaluating career development awards
    Survey of peer review strategies
    Repository of evaluations
- Networking and sharing ideas
  - e.g. Monthly & adhoc teleconferences
    Newsletters
    Annual meeting
    Partner exchanges





## Next steps for ICRP

- The database opens up new opportunities for individual and joint analysis & evaluation. Current areas of interest are: specific disease areas, environmental influences on cancer, prevention research.
- Expanding the dataset. We estimate that ICRP includes over 65% of world cancer research funding, but that still leaves a gap to fill
- Improving networking and increasing opportunities to share expertise
- Linking award data to research outcomes:
  - several pilot projects are underway in this arena
- Looking at mechanisms to make coding easier and maintain coding quality

# ICRP Partner experience with automated coding (Collexis/Elsevier)



# Background

CR-UK handles 700-900 applications for research funding per year. All applications coded to and

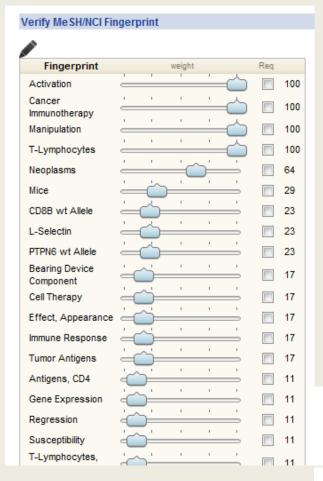
- **Type of Cancer:** our researchers code these. Quality is fairly good, with occasional 'hit the wrong button' errors
- Type of Research (CSO)
   We tried applicant coding (quality low), so we code these manually. This is time-consuming, so we initiated a project with Collexis to integrate automated coding into our grants management system.

#### Aims of the pilot:

- · Retain high quality coding
- Focus research managers' time on checking, not coding de novo
- Introduce a more reliable system



## System



Collexis' software generates MeSH keywords or 'fingerprints' of all grant abstracts.

The system gives us a list of keywords per abstract, which are ranked for relevance (we think using some kind of cluster/vector analysis based on the MeSH hierarchy).

This is used for peer reviewer finding, but also generates suggested CSO codes:





# Methodology

Collexis used a historical set of 3000 ready-CSO-coded CR-UK abstracts.

Collexis ran MeSH fingerprints on these and tried out a number of different methodologies to predict the CSO. The final mechanism used is a vector analysis which seemed to give good results on statistical analysis.

We waited for 7 months to accumulate a representative dataset across all areas. We are now in the process of evaluating the first 700 or so awards coded in this way.



# How is it working?

#### Initial results look encouraging

- Our first cross-check suggests that about 46% of the awards are coded to the same CSO sub-codes by the expert and automated coder and 92% have total or partial overlap between the major CSO codes applied.
- In the majority of cases, the automated system adds 1 or 2 extra codes (43%). We think that these additional codes are likely to be of relevance, but not major aims of the award, but some irrelevant codes are added.

Statistics - minor CSO code level	#	percentage
all grants	685	100.00%
# of grants where coding is identical between expert & autocode	21	3.07%
# of grants where all expert codes are picked up by autocode, but		
autocode adds extras	291	42.48%
# of grants where there is partial overlap of codes	204	29.78%
# of grants where no suggestions have been accepted	169	24.67%

However, 112 of these do agree at major CSO level (e.g. computer suggests CSO1.2, expert coder CSO1.3)



# Can we improve the algorithm?

- 1. Assess why the codes differ are there simple filters/keywords that could be suggested to improve quality:
  - e.g. One award investigating mechanisms of chemoresistance was expert-coded to 5.3, but coded to 3.3/4.3 by the algorithm. Can "chemoresist\*" be strongly associated to CSO5.3?
- Can we focus the algorithm better on the specific aims of the proposal? Will this help to remove 'extra codes'?
   e.g. In some instances, the extra codes applied by the computer are on peripheral concepts.
- 3. What's the comparison between expert-expert and expert-computer?

	Expert-exp	pert	Expert-Auto		
DRAFT analysis at MAJOR CSO level – 2007 study	#	%	#	%	
100% agreement	91	52%	208	30%	
100% & Partial agreement					
(same CSO major category or overlapping codes)*	161	92%	627	92%	
Completely different codes	14	8%	58	8%	

### What next?

#### For 2011 and first part of 2012

- We will continue to work with Collexis to improve the system, the system is designed to 'learn' and this is the first learning round
- If the results are good enough, we will stop manually coding unsuccessful applications

#### **Future:**

- Assess whether the auto-coding is fit for purpose for successful awards
  - if extra codes are added, does that matter?
  - is overlap at the major CSO category sufficient
- Is expert-expert variance similar to expert-computer variance?
  - large ICRP data validation study due to report January/February 2012

