

ChemBio Hub - capturing and sharing chemical biology information and knowhow



Brian Marsden
SGC
University of Oxford



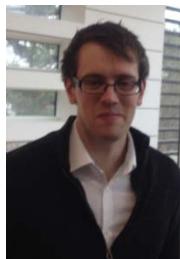
<http://chembiohub.ox.ac.uk>



Goals of ChemBio Hub

A place to share expertise, equipment, reagents, techniques and current areas of investigation associated with Chemical Biology.

- Improve internal communication
- Improve industry communication
- Encourage investment.
- Re-use/re-cycle pre-existing tools
- Make all available as Open Source





What is Chemical Biology ?

Human PIM1

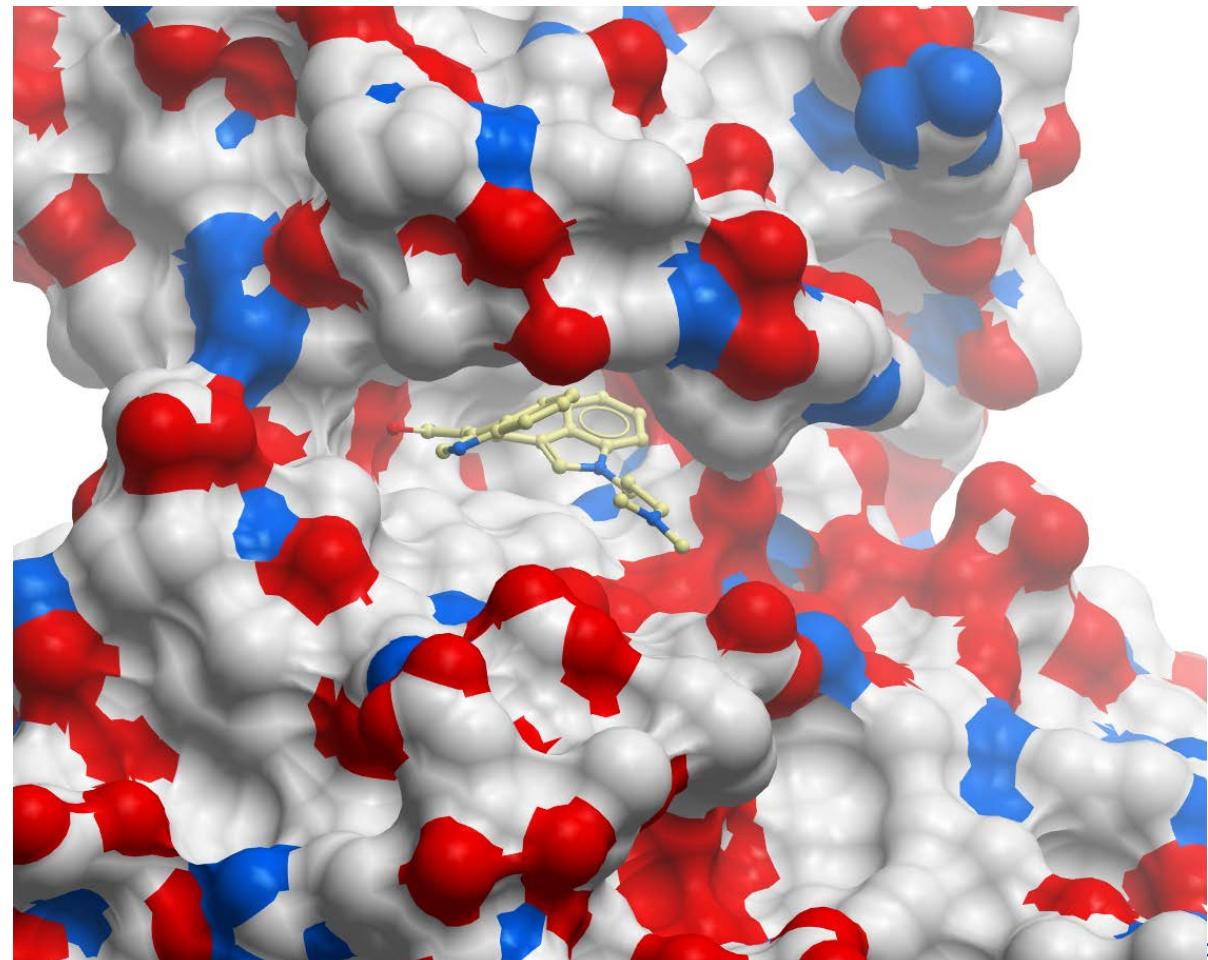
A protein kinase.

It puts a phosphate on other proteins as a means to cascade important signals within the cell.

When this goes wrong, cancer results.

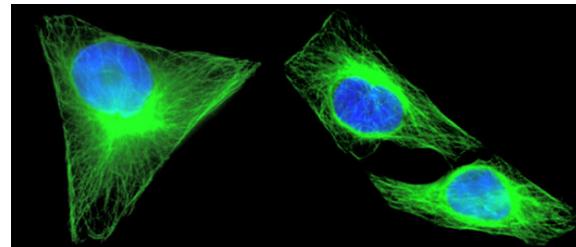
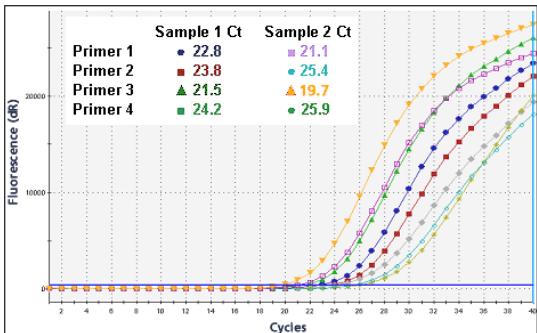
A small-molecule can help to prevent this by inhibiting the protein's default action.

But how do we interrogate this?





Generic assay (meta) data capture is very challenging



Immunostaining

qPCR data

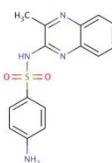
There is an incredibly diverse variety of assay data types



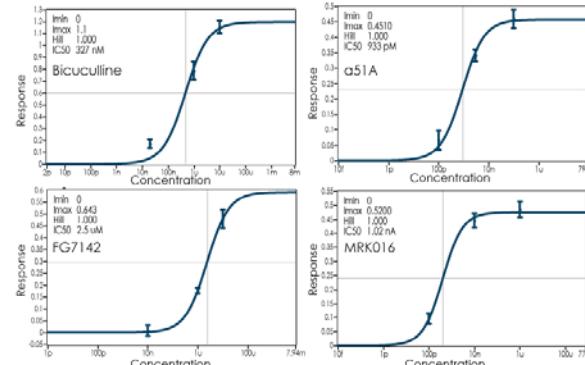
DMSO Control



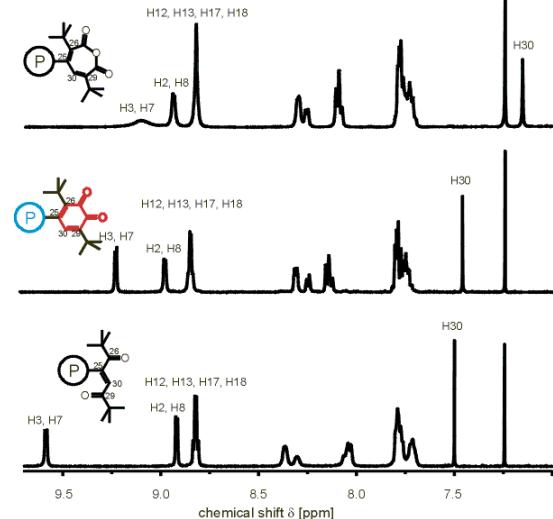
NCI 13156 C15H14N4O2S MW: 314.36



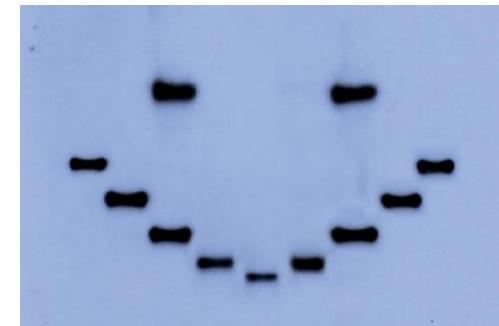
Phenotypic data



IC₅₀ data



NMR spectra



Western Blot



Challenges for University Researchers

- **Good science**

i.e. data management

- **Efficiency**

duplication of effort

- **Fulfilling Grant Conditions**

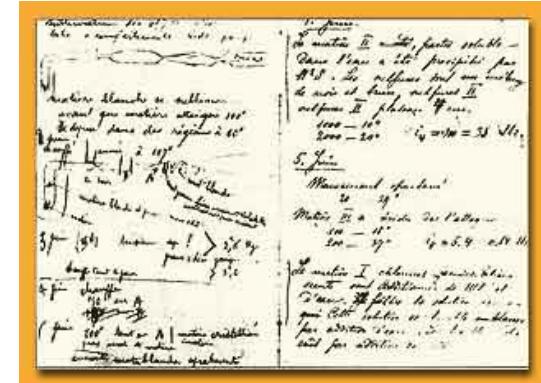
open data

- **Personal progression**

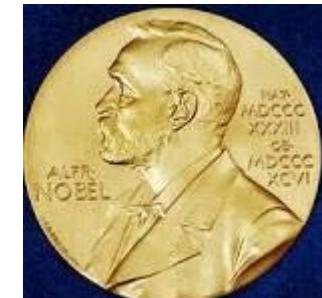
finding effective collaborators

- **Outcomes**

partnerships with industry to bridge the gap



wellcome trust



HELMHOLTZ
GEMEINSCHAFT

Open Science



ChemBio Hub – What the researchers say

We approached chemists and biologists direct and asked them about:

What they were using to store data (if anything)

What they disliked about their software

Whether sharing data was something they currently did

Lots don't use anything at all -
paper and pen and filed in a "database"...



Sharing was not widespread but scientists could see the benefit
Searching and better data processing were high priorities



Hasn't this already been fixed?

Sure, there are lots of tools available.





The solution – ChemBio Hub

Capture data, reagents/compounds, expertise –

- With assistance and curation
- In a central repository

Controlled levels of access –

- Within group, department, Oxford or externally

Example outcomes:

- The ‘go-to’ location for all aspects of University Chemical Biology
- Ability to discover tool compounds against target/protein of interest
- Identification of possible translational routes
- Pushing data externally, attracting pharma funding towards novel targets



The Web Platform

1.

Project List + Add Project

Filter project list: marsden project

Marsden Project

Draw compound Add multiple compounds

Search all compounds Search my compounds

Edit configuration Edit user roles Archive compounds

admin Restore / view archive



2.

Add Single Compound

If you want to add more than one compound at a time via SD or Excel, please use this form

Project Fields

Custom Identifier: Mars127

User defined ID

Name and Synonyms: 1 item selected.. BDM01

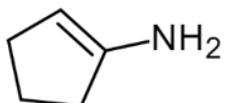
Compound aliases

Date of Synthesis: 19/05/16

Storage Conditions: 1 item selected.. Room Temp.

ChemDoodle[®]

Save compound Clear form



3.

Search

Marsden Project

Search ChemBio Hub Platform

Hidden Fields: Nothing Hidden

Sorts: No sorts

Filters: Upload ID: Pick From List: 2704

1 result

| ID in project | Structure | CBH_ID | Project | Custom Identifier | Name and Synonyms | Date of Synthesis | Storage Conditions |
|------------------|-------------|-------------|------------------|-------------------|-------------------|-------------------|--------------------|
| search-sort-hide | search-hide | search-hide | search-sort-hide | search-sort-hide | search-sort-hide | search-sort-hide | search-sort-hide |
| 1 | | CBHMT83M0Y | Marsden Project | Mars127 | BDM01 | 19/05/16 | Room Temp. |



Not just for chemistry...

Project List

Click a project title to see more details and add data to that project [i](#)

Filter project list: [i](#)

Demo InventoryReg Project



Inventory Data

- [Search all inventory items](#)
- [Search my inventory items](#)
- [Add single inventory items](#)
- [Add multiple inventory items](#)

Demo ChemReg Project



Chemical Data

- [Search all compounds](#)
- [Search my compounds](#)
- [Add multiple compounds](#)

Demo Compound Inventory



Chemical Data

- [Search all compounds](#)
- [Search my compounds](#)
- [Add multiple compounds](#)

Demo Antibody Project



Assay Data

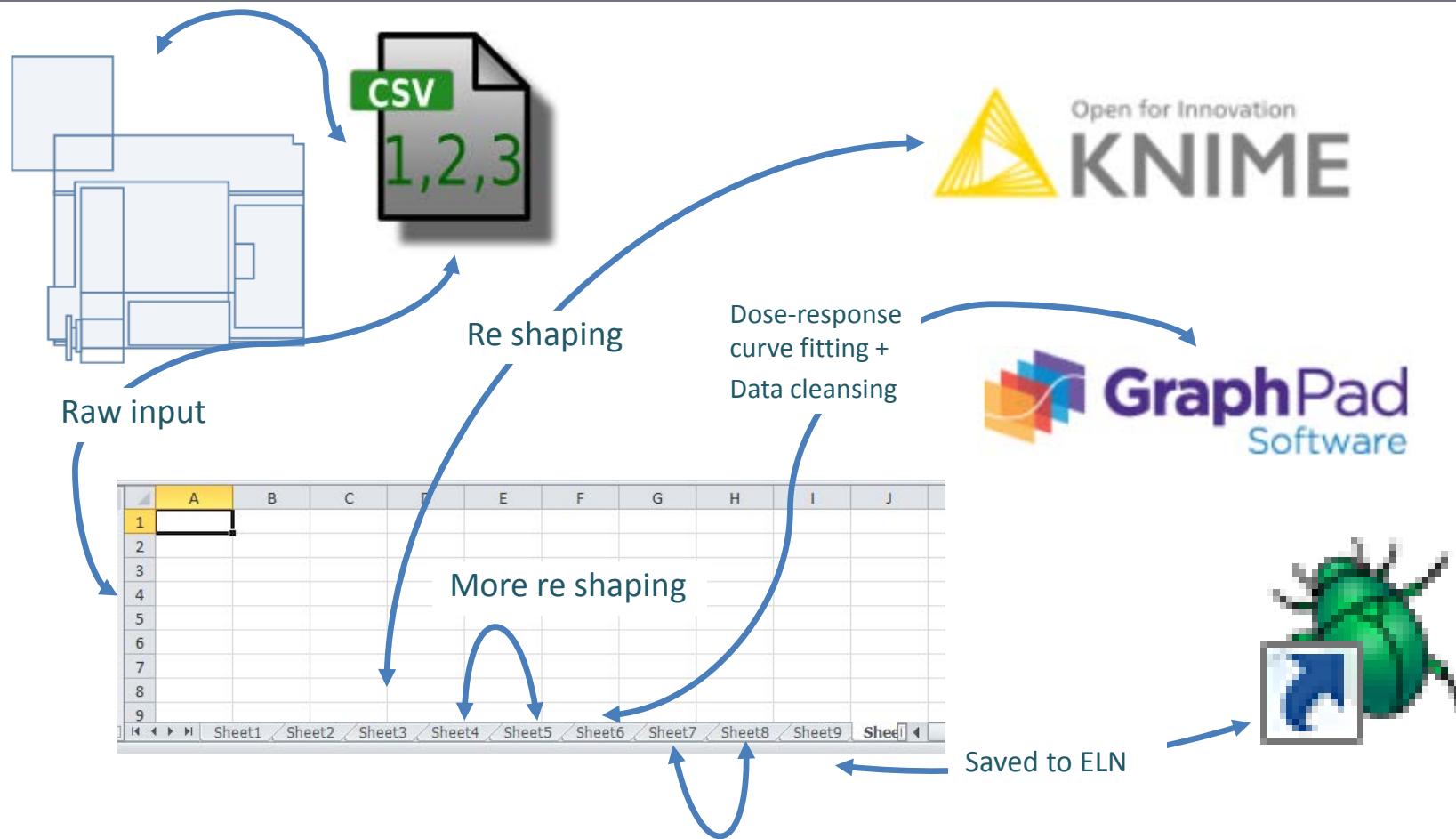
- [Data Overview](#)
- [Search My Data](#)

Inventory Data

- [Search all inventory items](#)
- [Search my inventory items](#)
- [Add single inventory items](#)
- [Add multiple inventory items](#)



Example Alpha Screen data flow





The solution: ChemBio Crunch

1) Upload raw data

Title*

BMG output file*
 No file selected.

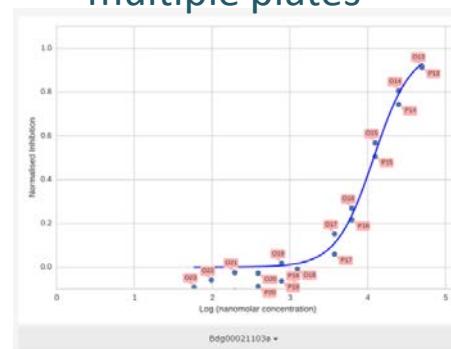
ESXX transfer file*
 No file selected.

Uploaded meta file
 No file selected.

2) Validate plates for systematic errors

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|--------|--------|---------|--------|--------|--------|--------|---------|---------|--------|--------|
| A | 600741 | 599507 | 932182 | 937140 | 932162 | 604212 | 939712 | 948132 | 932090 | 934546 | |
| B | | | | | | | | | | | |
| C | 56582 | 65588 | 86735 | 119054 | 115181 | 22689 | 951377 | 3226173 | 381292 | 938228 | 101463 |
| D | 52096 | 65075 | 85063 | 170326 | 149483 | 20025 | 20466 | 317623 | 3679515 | 932286 | 100163 |
| E | 294136 | 303952 | 415169 | 417122 | 319443 | 934603 | 552429 | 544713 | 599777 | 548777 | 531508 |
| F | 301777 | 390602 | 430188 | 830137 | 536791 | 546455 | 546700 | 545962 | 937137 | 399626 | 383649 |
| G | 51813 | 59964 | 70471 | 121289 | 129599 | 20875 | 271700 | 354906 | 354906 | 354906 | 504239 |
| H | 53529 | 62377 | 69369 | 172676 | 121592 | 511615 | 273611 | 357599 | 357599 | 473756 | 497403 |
| I | 87381 | 110390 | 155553 | 162000 | 28950 | 403440 | 490048 | 494048 | 939931 | 563037 | 568868 |
| J | 85129 | 109630 | 164767 | 170000 | 274664 | 400861 | 475000 | 492516 | 544063 | 599634 | 564718 |
| K | 82059 | 110295 | 140689 | 190001 | 100000 | 366333 | 416008 | 490018 | 500591 | 323562 | 540377 |
| L | 87533 | 110142 | 148131 | 160000 | 160000 | 361648 | 421228 | 494168 | 507141 | 323298 | 931551 |
| M | 47104 | 56343 | 88177 | 174196 | 186429 | 282493 | 346541 | 421018 | 474001 | 495159 | 704012 |
| N | 30931 | 51870 | 83505 | 124669 | 17529 | 287584 | 358644 | 401010 | 490594 | 493713 | 520291 |
| O | 378169 | 559360 | 4960335 | 577182 | 582616 | 564629 | 560220 | 574313 | 369494 | 872855 | 553394 |

3) Calculate IC_{50} for multiple plates



Mark as poor fit

| plate | compound | logIC50 | ic50 (nM) | system_comments | use | graph |
|-------|---------------|------------|------------|--|-----|-------|
| 2 | BDG00021 060a | 4.25927396 | 18166.6127 | Low total inhibition, values could be inaccurate | no | |
| 3 | BDG00021 344a | 4.8345796 | 68324.9943 | Low total inhibition, values could be inaccurate | no | |
| 4 | BDG00019 351a | 4.41049146 | 25733.0618 | Low total inhibition, values could be inaccurate | no | |

5) Export and deposit in ELN
(Comments automatically generated where IC_{50} may be inaccurate)



The technology stack

What you see



HTML



CSS



Bootstrap

ANGULARJS
by Google

What that runs on



WebAuth

Apache
HTTP SERVER PROJECT

PostgreSQL

ubuntu

What we use to deploy and test it



Travis CI



behave

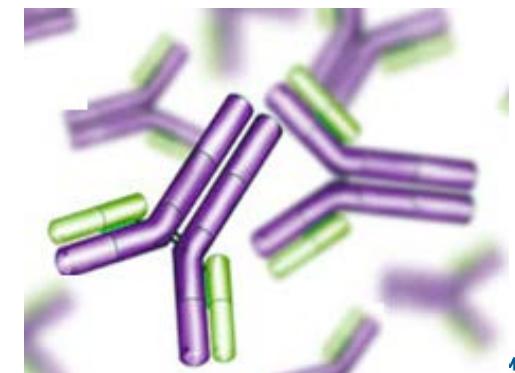
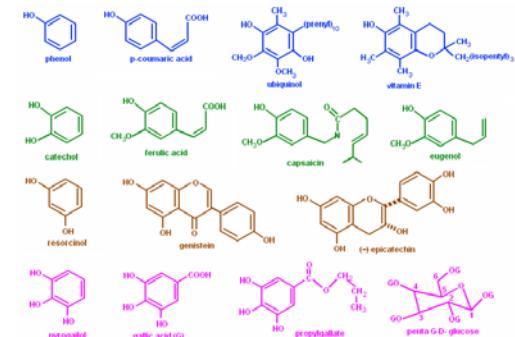




Who would use it?

Anyone who manages an inventory through a spreadsheet

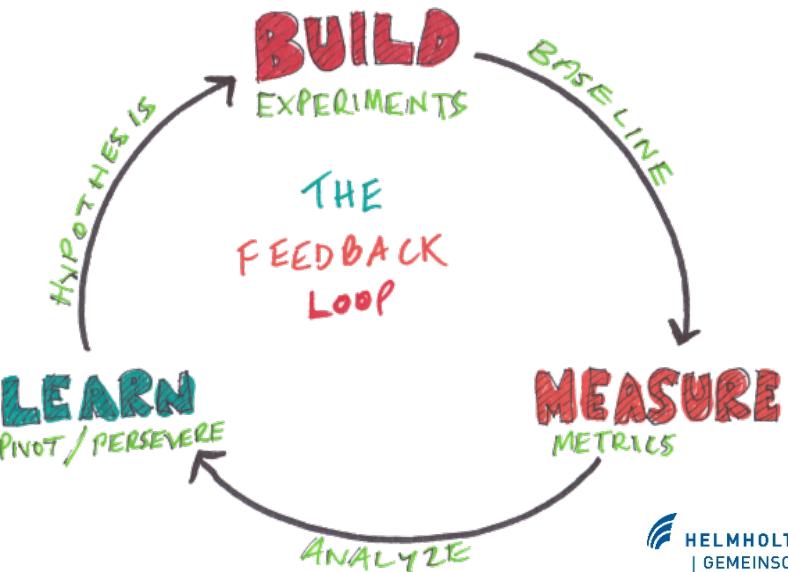
- Compounds
- Antibodies
- Cells
- Plasmids
- Lab orders
- Mice
- Unicorns





Why use it?

- Online
- Easy to use
- Search for things instantly
- Reduce duplication
- Free
- Academic project
- Evolving project
- Me!





Weaknesses of an evolving project

- Innate resistance to IT solutions
- System not quite up to scratch
- Unclear message
- Concerns about future
- Group engagement





Advantages of a data management system

For the PI -

Avoid the Post-Doc leaving panic

At the Bench -

Quickly search your data

Across the Department -

Save money – Sharing of resources can reduce departmental spending



Engagement

Engaged with 80 + labs across 10 departments:

- Chemistry
- Dunn school
- WIMM
- Pharmacology
- DPAG
- Plant Sciences
- Zoology
- Biochemistry
- Oncology
- TDI / SGC



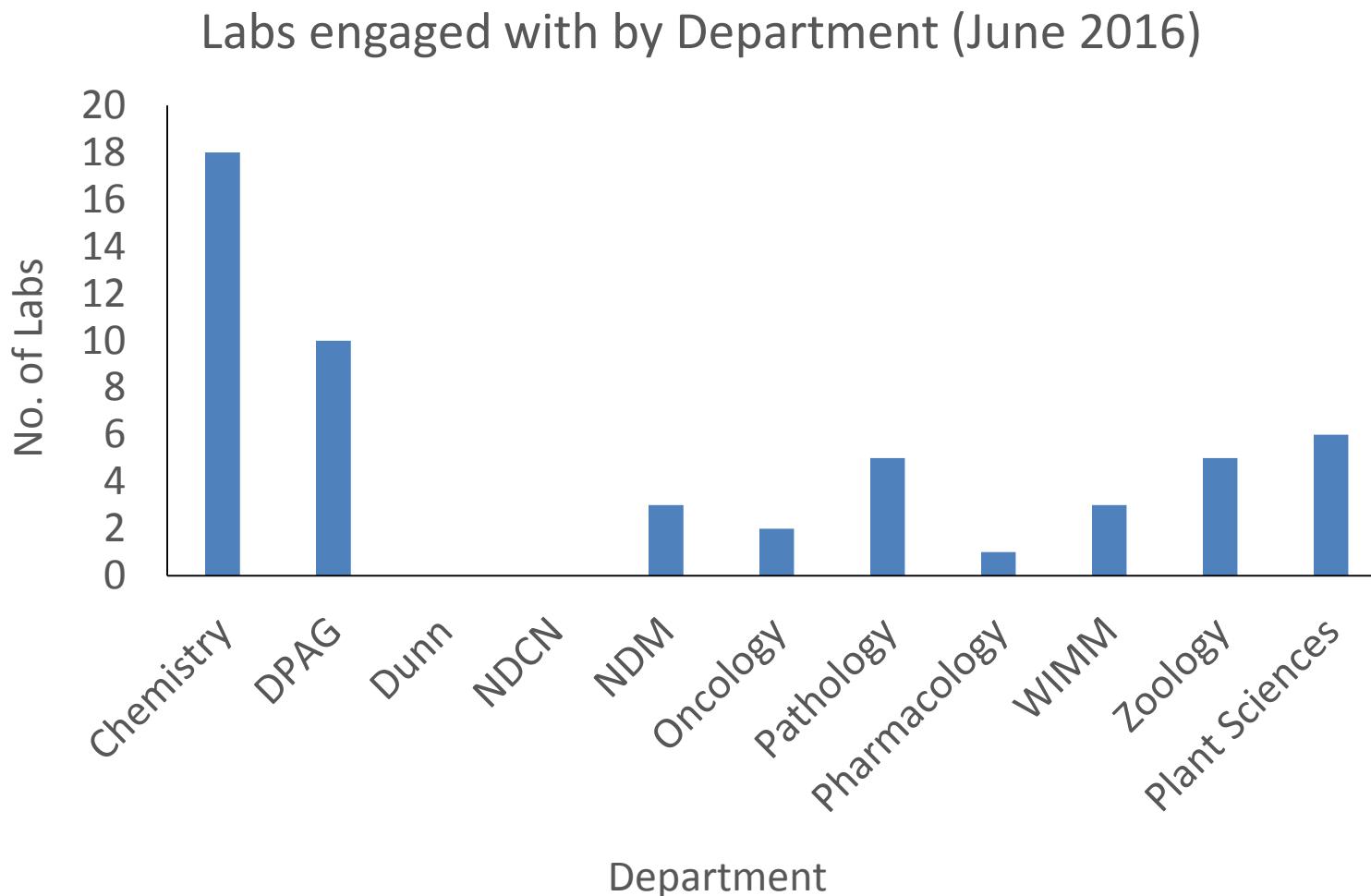
Engagement tools

- Emails / phone calls etc.
- Friends of friends
- ‘Dropping in’
- Departmental meetings
- ChemBio Hub symposium
- Trade stands



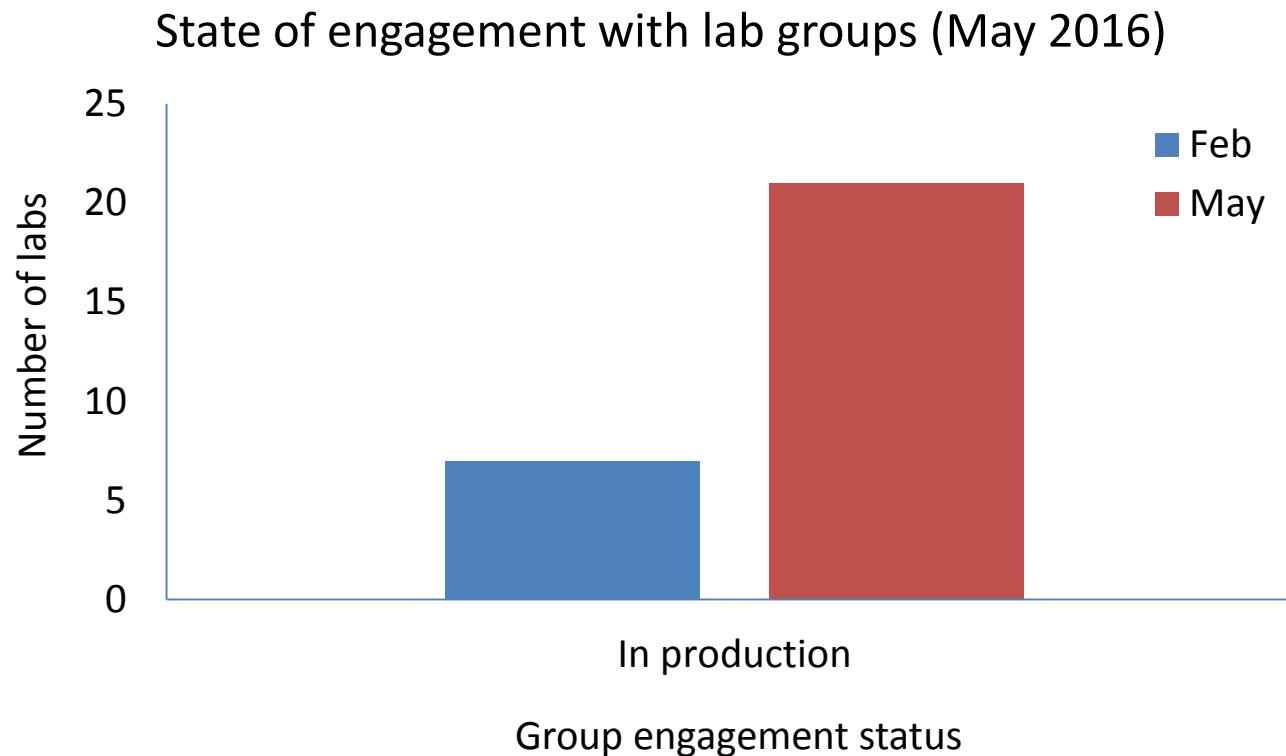


Engagement over time





Engagement over time





ChemiReg is in demand

- Clear systemic need for an Inventory manager
- People don't realise they have a problem
- IT support is not working as it should
- People bemused by computers



Clear University requirement for IT solutions

- Dunn School of Pathology stores inventory
- Chemistry departmental reagent inventory
- Chemistry Mass spec compound registration system
- Plasmid database system



Implementing change – Things to look out for

- Politics
- Novel solutions require an open mind
- Who is the lab authority?
- Lab morale can play a key factor



Things to keep on top of:

- Managing expectations is key
- You can't please everyone
- User aftercare
- Herding cats



ChemBio Hub



ChemBio Hub - capturing and sharing chemical biology information and knowhow

<http://chembiohub.ox.ac.uk>

Brian Marsden – Principal Investigator

Karen Porter – Project manager

Michael O'Hagan – KE coordinator

Andy Stretton – Developer

Paul Barrett – Developer

Adam Hendry – Project scientist

Prof. Alastair Buchan

Prof. Chris Schofield



Nuffield Department of Medicine
Medical Sciences Division



Supported by
wellcome trust



The John Fell
OUP Research Fund