Generating Solutions for Translational Research

Trent Schafer

Bachelor of Info Tech, Charles Darwin University

LabKey Database Developer

The Kirby Institute

Sydney, Australia

14th November 2011

Who is the Kirby Institute for Infection and Immunity in Society?

- University of New South Wales, Sydney, Australia

- Formerly the National Centre in HIV Epidemiology and Clinical Research

- Established in 1986 in Australia's response to HIV/AIDS

- Research   - diseases include HIV, viral hepatitis, STIs
                  - disciplines include laboratory science to population health

- 170 faculty and staff, including staff in Thailand and Cambodia.

Kirby Institute Research Programs

- Surveillance and Evaluation Program for Public Health

- Therapeutic and Vaccine Research

- HIV Epidemiology and Prevention

- Viral Hepatitis Clinical Research

- Viral Hepatitis Epidemiology and Prevention

- Sexual Health

- Aboriginal and Torres Strait Islander Health

- Biostatistics and Databases

- Immunovirology and Pathogenesis

- Research in the Asia-Pacific region

- Justice Health

- Public Health Interventions Research Group

Viral Hepatitis Clinical Research Program (VHCRP)

- Focus Areas:

  1. Co-ordinate multicentre clinical trials in HCV, HIV/HCV and HIV/HBV

  2. Conduct laboratory research in hepatitis C

  3. HepBank sample repository

- 20 FTE staff (4 academic, 12 general staff, 4 post-graduate students)

- Mix of public (NHMRC, NIH) and private (pharma) sector funding

- Sample repository

- Used by researchers world wide

- Custom built oracle database "HIVIRL"

- Researchers submit concept sheet for approval for sample use

- Process for use includes protocol steering committee approval, letter of understanding, ethics approval, etc

- Unused sub-study samples are returned to repository

- 780 subjects

- 5000 study time points

- 14610 samples
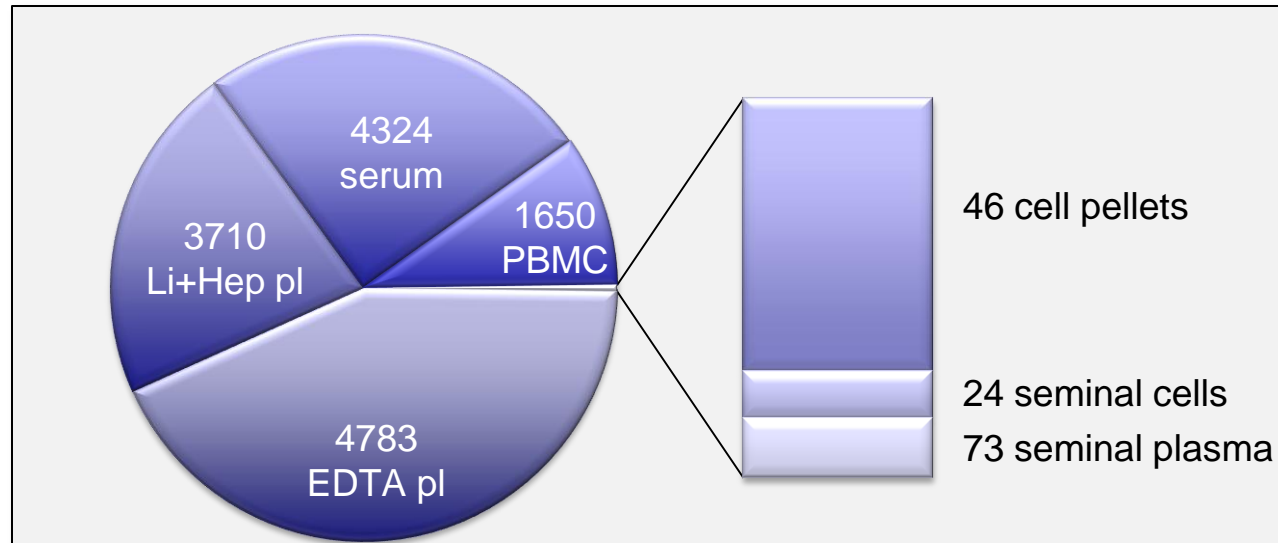
Sample Repository

			No. of patients
•	**ATAHC**	Prospective, natural history and treatment of aHCV	n = 163
•	**CHARIOT**	Phase IV induction dosing in cHCV	n = 896
•	**RAMPT-C**	Permucosal transmission in HIV+/- MSM	n = 32 / 45
•	**ITHACA**	ARYS / VIDUS, Vancouver longitudinal observational IDU cohorts	n = 500 / 1600
•	**ATAHC II**	Response guided treatment of recent HCV	n = 130
•	**DARE-C**	ATAHC II DAA substudy	n = 20, 4Q 2011
•	**ACTIVATE**	Phase IV, response guided, GT 2/ 3 cHCV active IDU	n = 100, 4Q 2011
•	**LAByRINTH**	Plasma Ribavirin prediction models	n = 20, 4Q 2011
•	**ETHOS**	HCV treatment in the opiate pharmacotherapy setting	n = 360, 1Q 2012
•	**SEARCH-C**	DAA resistance cohort in cHCV	n = 60, 1Q 2012
•	**TAILOR**	24 vs 48 wk PEG-IFN / RBV cHCV GT1 IL28B favourable, Asia	n = 400, 3Q 2012

What We Need

- Data consolidation

- Improved search ability across clinical, and laboratory data

- Specimen tracking system

- Audit trail

What We Have and What We Need
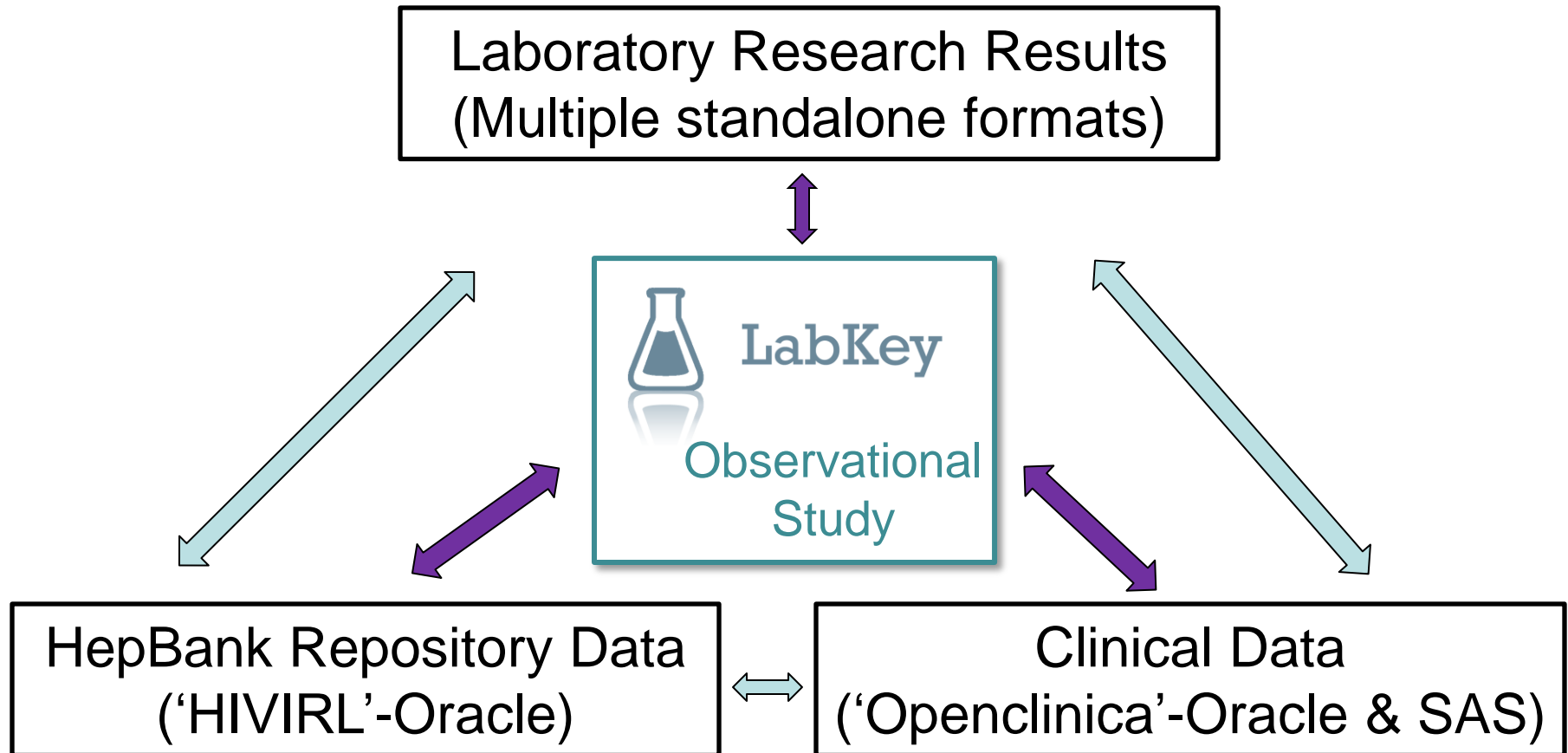
What LabKey Can Provide

Free, open access internet-based data entry and searching

- Clinical, specimen & research data linkage

- Secure storage, permission levels

- Sample searching, requests & approval process

- Longitudinal results, audit trail

Facilitate manage & share lab data

- Raw data and compressed results

- SOPs, experiment & equipment details

- Minimum data & reporting standards

Custom Enhancements

1. Oracle Database

2. Time chart for visits

3. External repository

4. Bulk sequencing

1. Oracle Database: Requirements

- Ability to create an oracle data source

- Export request per mL

- Up to date information

1. Oracle Database: Design

- Requested per mL rather than quantity

- Live view of Sample Repository

- Text box for quantity

- Export Data for updating HepBank

1. Oracle Database: Design

Step 1. Select vials and click start request

Step 2. Specify desired volume and export

1. Oracle Database: Decision

- Re-inventing the wheel

- All vials returned to repository

- Manually create spreadsheet

1. Oracle Database: Implementation

- Oracle user with read access to sample repository

- 2 custom oracle packages – zip, exporter

- Windows share to write repository to

- Scheduled to run daily

2. Time Chart for Visits: Requirements

- Using time points instead of date visits

- No built in controls for creating time charts on visits

- Use Google charting API, LABKEY.Query API and a custom SQL query to produce a time chart with time points.

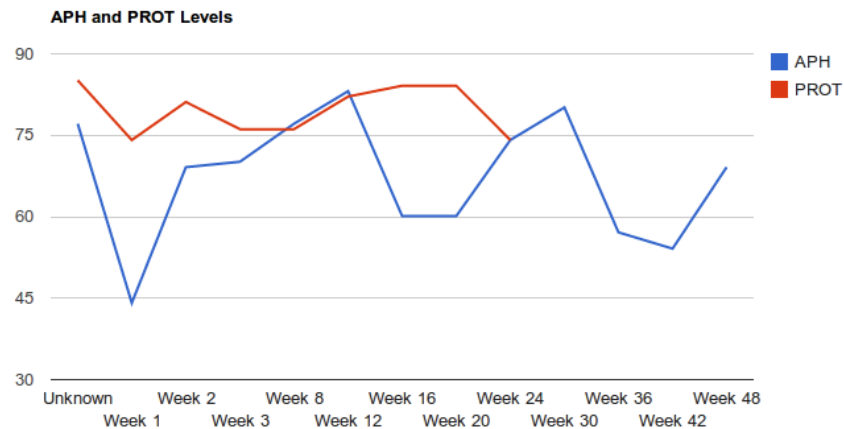# 2. Time Chart for Visits: Process

- Step 1. Create a Query in the schema browser
- Step 2. Customize participant view to draw a chart

3. External Repository: Requirements

- Capture new specimens coming in (from external sources)

- Tracking so can determine if specimen is received or not

- Export received specimens in batches

- Ability to import into custom built specimen repository

3. External Repository: Design

- Separate project folder for each external group due to permissions

- Two folders

  - Submission folder: list with list lookups

  - Approval folder: list linking to submissions

- Series of custom SQL queries, joining from the two folders for nice reporting

  - received, not received

4. External Repository: Design

Step 1. Create a patient visit

Patient Visit

Patient ID	111
Namecode	ABCD
Week Visit	2
DOV	2011-10-10

Save Visit Clear Form

DISCUSS THIS ▸

Alert ✕
Successfully Saved Patient Visit
OK

Step 2. Create a vial which is associated to a patient visit

Namecode	ABCD
Week Visit	2
DOV	2011-09-10

Save Visit Clear Form

Visit Vial

Vial Type	serum ▾
Box Name	rrrr
Box	1
Location	1
Volume	1
Spot ID	1
Exponent	1
Removed	1
Comments	comments

Add Vial Clear Form

Alert ✕
Successfully Added new Vial
OK

4. External Repository: Design

Step 3. Receiver marks a vial as having been received

Un-Received Specimens

VIEWS ▼ QUERY ▼ EXPORT ▼ PRINT PAGE SIZE ▼ [MARK RECEIVED]

	Patient Visit Vial Id	Patient Visit ID	Vial Type	Box Name	Box	Location	Volume	Spot Id	Comments	Exponent	Removed	Namecode	Week Visit	Dov	Patient Id
☑	7	111	serum	rrrr	1	1	1.0	1	comments	1	1	ABCD	2	2011-10-10	111
☐	8	111	serum	gggg	2	2	2.0	2	sdfsdf	2	2	ABCD	2	2011-10-10	111

Received Specimens

QUERY ▼ VIEWS ▼ EXPORT ▼ PRINT PAGE SIZE ▼

Patient Visit Vial Id	Patient Visit ID	Vial Type	Box Name	Box	Location	Volume	Spot Id	Comments	Exponent	Removed	Namecode	Week Visit	Dov	Patient Id	Date Received
4	154	serum	FFF	1	1	1.0	1	dfdf	1	1	DDDD	3	2011-10-03	154	2011-10-10
5	987	serum	ABC	1	82	0.5		Thanks Trent			ZYBS	3	2011-10-09	987	2011-01-01
6	12654	serum	NAH	1	58	0.5					JDIW	4	2011-10-09	12654	2011-10-10

3. External Repository: Queries

- All vials with status

```sql
select av.*,
case
        when vr.vial_id is not null then 'RECEIVED'
        else 'NOT RECEIVED' end received
from all_vials av
left outer join "/Home"."Offsite Repository"
."AMRv2 Specimen Approvals"."lists".vials_received vr
on av.patient_visit_vial_id = vr.vial_id
```

4. Bulk Sequencing: Requirements

- Machine used is a 3730xl DNA Analyzer by Applied Biosystems, and produces AB1 sequence files

- Point to multiple files to run the insert in one go

- Ability to load in those files and produce chromatogram

4. Bulk Sequencing: Design

- Using third party library to extract sequencing data

  - BioJava is an open source library with support for multitude of biological data http://biojava.org/wiki/Main_Page

  - Legacy version has an AB1 parser

- Initially – Export to TSV. Import into LabKey

- Visualize the Chromatogram

- Future – Hopefully streamline the process so the above operation is just a matter of uploading a file

4. Bulk Sequencing: Code Samples

- Using same line chart as in Time Charts
- Example chromatogram:
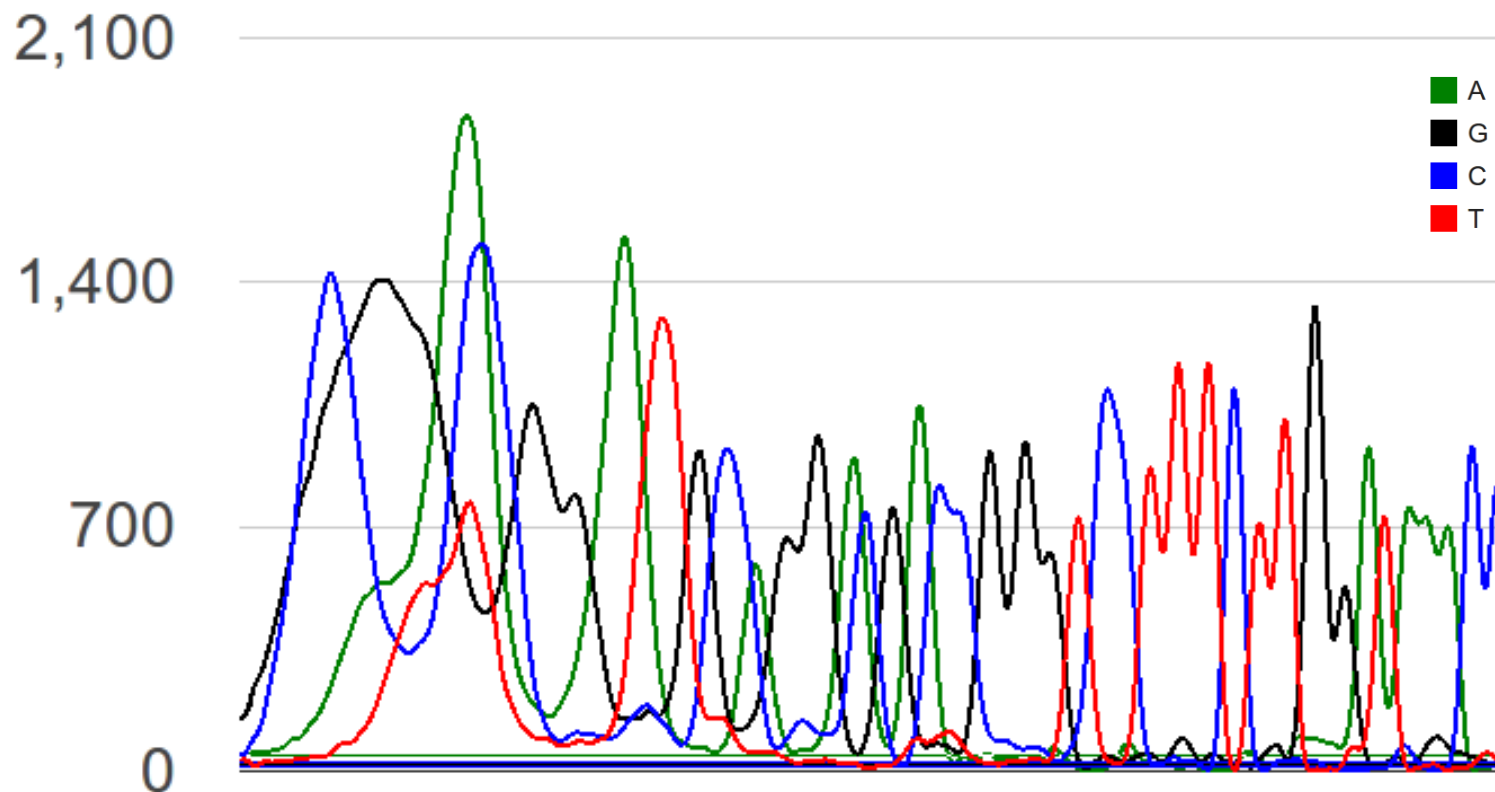
Thank You

- Thanks for listening

- Thanks to University of NSW for funding this work (Major Research Equipment and Infrastructure Initiative [MREII])

- Thanks to Adam Rauch for helping with the Oracle data source

- Thanks to all who have helped me in the forums

- Thanks to the wonderful team I work with