



**Data Network Expert Round Table Summary
Monday 16th June 2003
Canberra – ABS House**

ATTENDANCE (TABLE ORDER – ANTI CLOCKWISE)

Main Table

- Fiona Stanley (Alliance)
- David Johnson (Melb Institute)
- Robert Wooding (Health & Ageing)
- Alan Pettigrew (NHMRC)
- Richard Madden (AIHW)
- David Hazelhurst (FACS)
- Karen Wilson (FACS)
- Diana Hetzel (Uni Adelaide)
- Brian Oldenburg (Public Hlth QUT)
- Grant Tidswell (Centrelink)
- David De Vaus (AIFS)
- Lawrence Cram (ARC)
- Brian Richards (HIC)
- Dave Abel (CSIRO)
- John Bass (Curtin Uni)
- Robyn Sheen (Prod Comm)

- Dianne Heriot (Attorn General)
- Susan Linacre (ABS)
- Richard Bridge (DEST)
- Ching Choi (AIHW)
- Helen Moyle (AIHW)
- Tony McMichael (NCEPH, ANU)
- Michelle Fisher (Priv Comm Vic)
- Jonathan Palmer (ABS)
- Geoff Rasmussen (Facilitator)

Side Table

- Lynne McGuigan (Alliance)
- Carey Drake-Brockman (Alliance)
- Alan MacKay (ABS)
- Brant Pridmore (Federal Privacy Comm)

SESSION ONE – SETTING THE SCENE

1. Introduction and Benefits to Australia

Fiona Stanley started the day by introducing the notion of a national effort on better data sharing and data linkage. She spoke to the generic benefits this would bring to Australia in terms of both research and policy/practise. Subsequently, Dave Abel and Jonathan Palmer confirmed the strong support of CSIRO and ABS to this topic and discussed how it fitted into their agencies strategies and business plans.

Each workshop participant spoke for 30 seconds (or more) to introduce themselves and to summarise their view on the benefits and/or the key challenges presented by this topic.

Fiona also summarised the WA experience and, in particular, talked to the research and intervention assessment the linked data network had enabled in WA in recent years.

2. Round Table: What requirements would the data network need to meet to secure the support of those present

This session was held as a brainstorming discussion. Items were collected under two headings: (1) Outcomes Targeted; and (2) Style/Constraints within which the Data Network would operate. The items raised are listed below. Some items ended up being suggestions for near term actions and these have been grouped together under a third heading. Finally, several requirements for success came up in the Agency Position Statements that were provided prior to the workshop. These are also summarised below under a fourth heading:

Group	Issue
Outcomes Targeted (from 16 June)	<ul style="list-style-type: none"> ▪ Clearly understood purpose of the network (is it linked up data or joined up data?) – G2 ▪ Integration of national agencies – G2 ▪ Capacity within Australia to interpret linked data would need to be improved (includes funding for these endeavours)... related discussion around “Enabling Grants” from NHMRC and “Network Grants” from ARC – G2 ▪ Network would be Issue Driven rather than Data Driven – G1 ▪ Cross-jurisdictional insights and conclusions should be possible from data linkages – G2 ▪ The network would become a resource for monitoring and evaluating policy and interventions (eg. impact of Folate program in WA) – G2 ▪ Saves time for researchers – ease of access to data; ease of comparison of data ▪ Network enables advocacy for filling key gaps in Australian data and advocates improvements in data quality ▪ In particular, advocacy includes focus on gaps in non-administrative data sources (eg. Pre-school records) ▪ Supports awareness of privacy issues – G1 ▪ Access to data is improved – as simple as knowing what is available and how to get it ▪ Provides access to micro-population issues through mass population scope (cf Surveys) – G2
Style & Constraints within which Network should work (from 16 June)	<ul style="list-style-type: none"> ▪ Clear focus on improving consistency and quality of data ▪ Clear procedure for engagement of State Govt and Local Govt regarding their data – G1 ▪ Compliance with privacy legislation (state level; federal level) – G1 ▪ May require changes to national privacy regime and ethical clearance structures? – G1 ▪ Very strong and transparent governance – G1 ▪ Transparency of ownership issues ▪ Purpose clearly known when data is accessed and permissions linked to this purpose – G2 ▪ Manage risks of misleading data – G2 ▪ Manage risk and sensitivity to data being used for public debate/mass media “stories” rather than research/policy/practise – G1 ▪ Availability of data linked to purpose – not necessarily always available – G1 & G2 ▪ Sensible progression – small steps to longer term goal – don’t bite off the whole thing at once ▪ Sound de-identification process and management of re-identification risk – G1 ▪ Finds balance between data utilisation value to research/policy/practice vs re-identification risk – G1/G2 ▪ Don’t sacrifice primary purpose of data collection efforts to achieve secondary goals associated with linkage – G1

Group	Issue
Suggested Action Items (from 16 June)	<ul style="list-style-type: none"> ▪ Potential demonstration project – target a change to the 2006 census to achieve UK style retention of 10% records for subsequent linkage efforts ▪ Need to promote the benefits/value of data linkage to respond to the concerns raised about privacy ▪ Should promote national data dictionaries (knowledge and compliance) so as to ease future linkage efforts ▪ A mapping exercise is needed to stocktake what data we have and what is in the pipeline ▪ Should provide clear summary of achievements in WA through data linkage as selling tool (later added UK/Scandinavia to this)
Issues Raised (from Agency Position Statements)	<ul style="list-style-type: none"> ▪ Purpose of Use – G2 ▪ Permissions – G2 ▪ Ethics – G1 ▪ Confidentiality – G1 ▪ Data Standards ▪ Linkage ▪ Network Reporting ▪ Citation & Acknowledgment – G2 ▪ Resources for Collection – G1 ▪ Security – G1 ▪ Privacy – G1 ▪ Licensing ▪ Costs of establishing & maintaining ▪ Legislative compliance – G1 ▪ Policy/framework compliance – G1 ▪ Data collection focus – G1 ▪ Quality ▪ Access Regime – G2 ▪ Intellectual property rights ▪ Infrastructure ▪ Billing? ▪ Checking – relevance/up-to-date/accuracy & reliability ▪ Source identification ▪ Consultancy & risk of misinterpretation – G2 ▪ Costs to contribute ▪ Encouraging contributions

Table One – Requirements for Success

SESSION TWO & THREE – ISSUE DEBATE

In preparation for the second session, the group was split into three sub-groups. These were to target:

- ¶ Group One – Issues associated with contribution of data to the network (including the G1 issues above)
- ¶ Group Two – Issues associated with the utilisation of the data in the network (including the G2 issues above)
- ¶ Group Three – General strategy and purpose of the data network

1. Results from Group One

The group focused on the governance model that should be used for the data network. They used the slide below to summarise their thoughts:

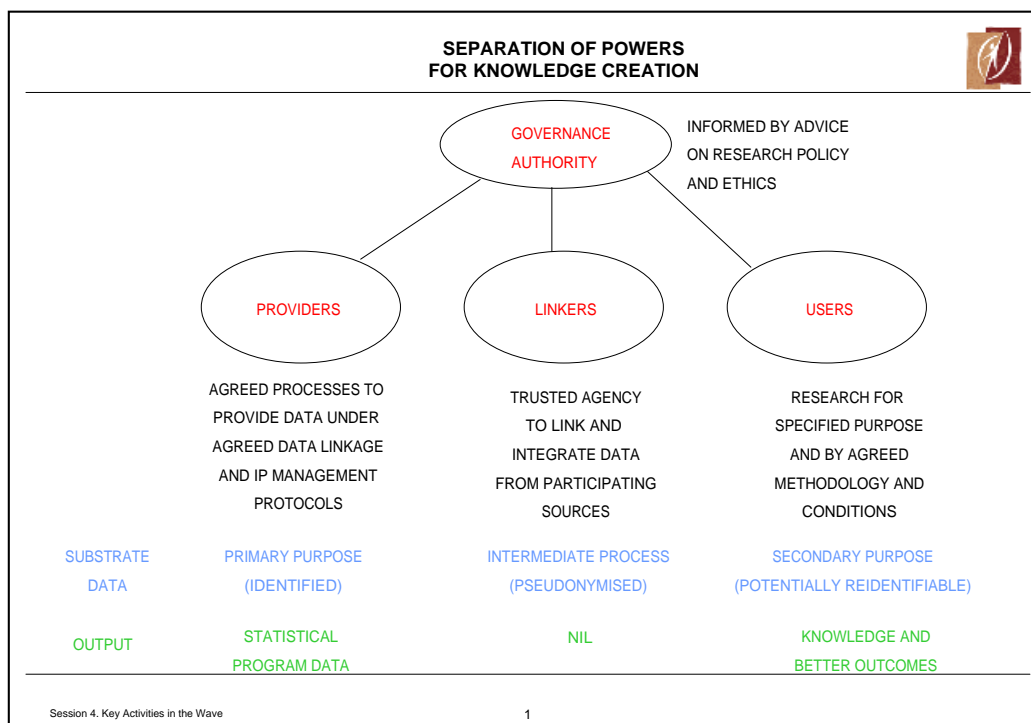


Figure One – Group One’s Suggested Governance Model

Results from Group One (cont)

Topic	Issues documented during the breakout
Issues	<p>Communication regarding the Data Network needs to be underpinned by clear value propositions regarding its role:</p> <ul style="list-style-type: none"> ▪ Savings on supply side re data ▪ Outcomes and benefits to the community ▪ Providers of data – better understanding of role – see data contribute to a greater purpose <p>Need for a highly ethical governance mechanism to reassure data providers.</p> <ul style="list-style-type: none"> ▪ Cost to providers of supplying data that is suitable to be interrogated by others ▪ Data to be linked would be determined by the research questions – NOT – if we build it – they will come! ▪ Governance and protocols are the key issues not the technical side
Proposed Governance	<ul style="list-style-type: none"> ▪ A ‘Ghandi’ like body – inspiring respect and confidence ▪ Prefer one ethics committee to reduce bureaucratic steps ▪ Voice of consumers must be represented ▪ Need to separate out some of the roles in the process: determining if research should proceed (ethical issues), funding of research, approval of linkage etc.
Key Challenges	<ul style="list-style-type: none"> ▪ Loss of control by governments will be an issue – protection from political interference ▪ Obtaining the necessary level of co-operation. Trust in protocols will be required. ▪ Re privacy – the political and public positioning responsibilities are as important as the legal ones.
Suggested Directions	<ul style="list-style-type: none"> ▪ Health Consumers Council are interested in having a discussion paper on these issues ▪ Commonwealth would need to fund a data linkage unit – ARC and NHMRC jointly?

Table Two - Results from Group One

2. Results from Group Two

Topic	Issues documented during the breakout
Purpose	<ul style="list-style-type: none"> ▪ Primary purpose – national capacity to use data more effectively ▪ Purposes – clarity and availability... with permissions linked to purpose ▪ Joined up data ▪ Researchers will use the best possible sources of data ▪ Impact of broader social and economic change on child outcome indicators – need for much broader info eg work place changes (Bronfenbrennar ecological model) ▪ Ability to map impact of large scale events on a population group ▪ Administrative data – look at broader impacts of agency functions
Overcoming Hurdles	<ul style="list-style-type: none"> ▪ Sensitivities of government agencies and politicians to release of info – can also build relationships with Ministers ▪ 96% agreement to linkage from WA aboriginal study participants – maybe privacy not such a huge issue ▪ Ongoing state based linkage can be quite immediate ▪ Utilisation of the data at the local level – also needs funding and development of new models for dissemination ▪ Sensitivity around notions of a unique identifier – trust takes time to build – convey message about the benefits that can derive from information. Possible role for a non-government leader in re-opening the debate. Demonstration projects can help. System needs to proceed and not be dependent on this eventuating ▪ Control over outcome reporting... surprises are eliminated if key people are involved in the decision making. Should not be researcher only driven
Key Challenges	<ul style="list-style-type: none"> ▪ Where are the gaps – lots of data – pulling it together the problem ▪ National collections slow and delayed ▪ Beware about increasing complexity by imposing a model – can recommend practices – process will be iterative. Not all research is directly policy relevant but may still be important. Focus may be cross-disciplinary collaboration, but intensive within-silo work not precluded. Open fishing expeditions can undermine trust and reduce willingness to share. ▪ Academic competition can undermine collaboration – but children and young people a more co-operative sector. The way that research is rewarded and funded maintains this problem.
Suggested Directions	<ul style="list-style-type: none"> ▪ Address gaps in funding of existing data bases – build on available strengths ▪ Look at multiple functions that data could serve – not just restricted to the collecting agency's purposes ▪ Value of beginning to build some 'shared data' infrastructure so that each individual project does not have to reinvent the wheel individually – could spend more time on analysis rather than linkage ▪ Is the resource we need a gatekeeper role? Clearinghouse function. The infrastructure can build slowly in relation to the specific research questions. A facilitator and enabler and an advocate for the national benefits. ▪ Potential to undertake an ethics monitoring role as well. Data collecting bodies all have their own mechanisms but linkage mechanism will need it also ▪ Broker role that harnesses the information that is available. Needs skills in social epidemiology and relationship building

Topic	Issues documented during the breakout
Recommended Way Forward (as presented back to the whole group)	<ul style="list-style-type: none"> ▪ Be clear about purpose – improving the wellbeing of children and youth ▪ To do this need broad multidisciplinary approach – key central issues and broader socio-economic factors in the context (ecology) ▪ Research and policy people need access to data to answer key questions – build an infrastructure to serve this agenda <ul style="list-style-type: none"> – Linkage – Provenance of the data, ethics, control over access, authentication, protocols, management protocols for use and access to the data, broker role ▪ How we take the process forward <ul style="list-style-type: none"> – Build on what has already been achieved – Learn from international experience – Models of regional data to enable policy and practice agendas – contrast between states (Each state has strengths and weaknesses – different legislation etc) – Examples of building up trust – Small pilot projects – both local level and broader – Model of HIV/aids

Table Three - Results from Group Two

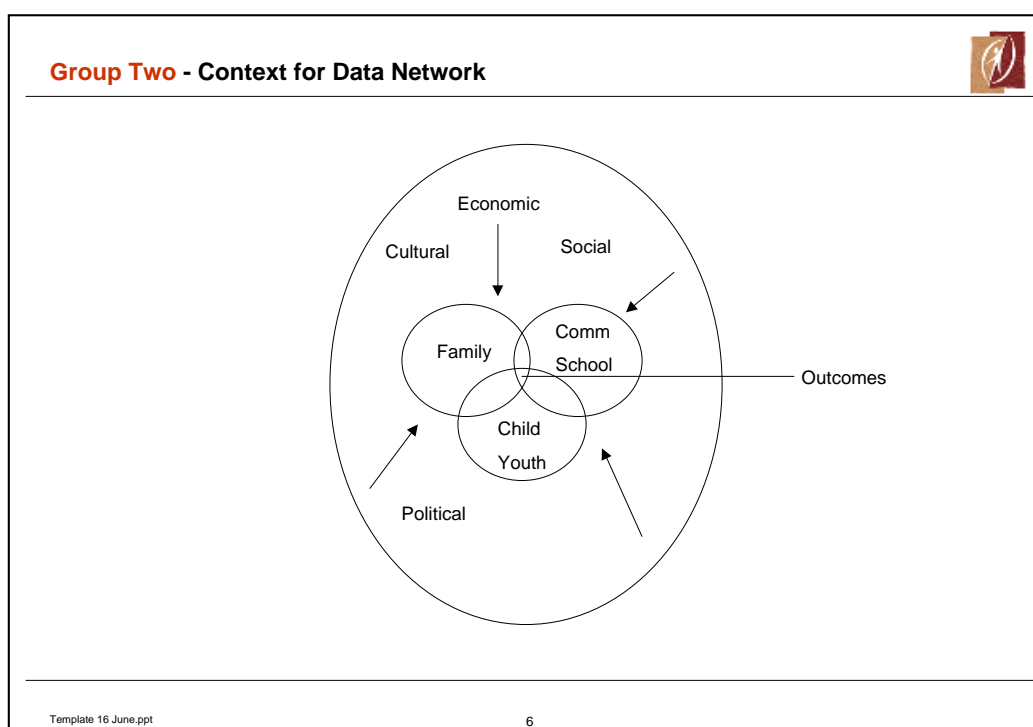


Figure Two – Group Two's Context Explanation

3. Results from Group Three

Topic	Issues documented during the breakout
<p>Recommended Way Forward (as presented back to the whole group)</p>	<ul style="list-style-type: none"> ▪ Shift in research and funding focus - whole of life research, gaining insights into complex problems, need to get serious about longitudinal research and facilitate upstream causes and diverse outcomes. To do this need linked data that go well beyond health and welfare. ▪ Agree that Alliance should be a virtual network. Not a physical data warehouse. ▪ The data network is also the linking of people ▪ Market research around the barriers to data sharing – this is something Alliance could do ▪ Focus on definitions and standards will facilitate culture for sharing data ▪ Mapping of data sets that currently exist – and who is using them ▪ Produce a description of use of linked data and the resulting scientific benefits to come from it. This would be would be a helpful marketing tool. It would show the value of linking up data sets. Not cheap ▪ Extending the principles for the use of data that has been well set in health, across other sectors. This would ensure that at least across the Alliance we have an agreed set of principles. Coalition of data custodians that are part of the Alliance will want to know that these exist. ▪ Develop a prototype that can demonstrate value and issues of joined up data and market data sharing – and to push the boundaries of how researchers work ▪ Alliance, possibly, to secure the funding for a Conference of data experts to facilitate networking and build support

Table Four – Results from Group Three

SESSION FOUR – TAKING THE DATA NETWORK FORWARD

We started this section by discussing a framework for thinking about the Data Network. Essentially, it has three goals:

1. To improve **access** to and **sharing** of data available for research, policy and practice
2. To introduce and improve **data linkages** enabling better research and better evaluation of policy and practice
3. To improve **completeness** of data to facilitate goals 1 & 2 above

On each of these three goals, the risk is that we try to go too far too fast and subsequently achieve very little. There was strong support from within the group for a pragmatic, staged approach to implementation of the data network. Equally, we need to be driven by the vision of where this is going, to ensure that our near term actions are consistent with this vision. Figure Three, below, summarises the framework that resulted. We have then used its components as headings for what work should be done going forward – see Table Four. Figure Four summarises many (but not all) of the suggested projects against the framework. Table Five, however, is the comprehensive record of what was agreed as next steps together with volunteers to run with these steps.

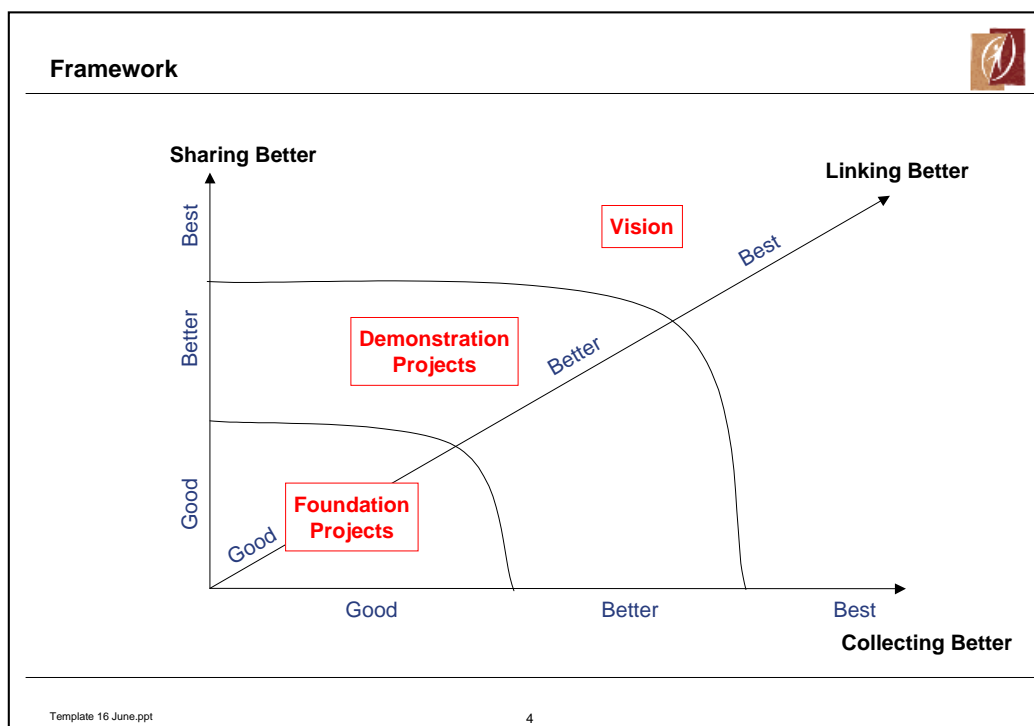


Figure Three – Framework for taking Data Network Forward

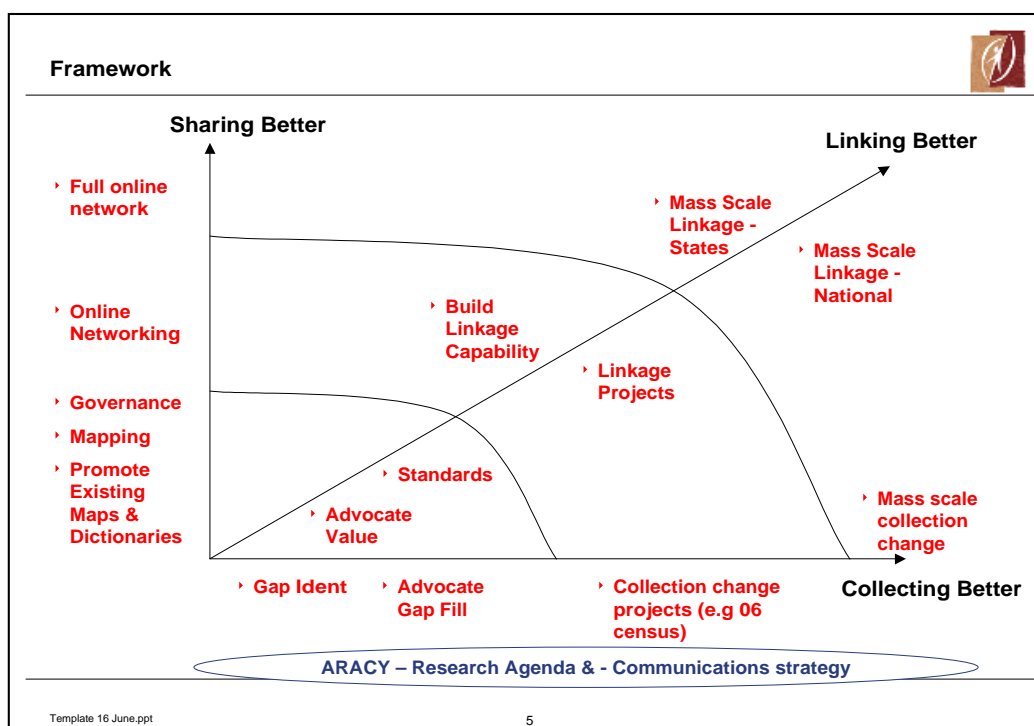


Figure Four – Framework Partially Filled Out

Dimension	Maturity Stage	Action	Volunteers
Data Sharing	Foundation Projects	1. Organise conference on data sharing (includes leverage into policy and practise)	L Cram
		2. Develop governance model based on suggestion presented today	S Linacre/R Wooding
		3. Define distributed architecture; define architecture for networked meta data	J Palmer
		4. Advocacy for better adoption of data standards (primarily targeted towards researchers)	C Choi/R Bridge
		5. Legal guidelines/framework for sharing and linkage	S Linacre/M Fisher
		6. ARACY data brokerage service (not just now)	B Oldenburg
	Demonstration Projects	7. Map of existing data efforts (limited scope initially)	A Mackay
		8. Map of meta data and standards in use	C Choi/A Mackay
		9. Pilot online network and grow organically	D Abel
		10. Children with disabilities	H Moyle
		11. Indigenous children	S Zubrick (Fiona to discuss)

Dimension	Maturity Stage	Action	Volunteers
Data Linkage	Foundation Projects	12. Skill/capability building effort (associated with building and using linked data sets) – includes undergrad/postgrad programs; proposal to ARC/NHMRC for funding of capability building effort	L Cram
		13. Governance model (see 2 above)	See 2
		14. Document value potential from linkage using WA and overseas examples... turn into marketing document for benefits of data network	Alliance
		15. Survey of ethics committees positions on data linkage	R Madden
		16. Legal guidelines (see 5 above)	See 5
		17. Alliance funding submission for network development	Alliance/T McMichael
	Demonstration Projects	18. Immunisation WA/SA project	J Bass
		19. Child protection/education link project	H Moyle
		20. Tasmanian Birth – Hospital – ATO – Centrelink link project	G Tidswell
		21. Immunisation compliance causes and effects	J Bass
		22. Local populous bottom up design for data required to assess local interventions (EDI link)	D Hazlehurst/F Oberklaid
		23. Alliance research agenda priority projects	Alliance
	Gap Filling	Foundation Projects	24. Advocacy for under resourced databases (includes resource solution for birth defects database; includes APSU)
25. ARACY research agenda (should drive data collection efforts)			Alliance
26. Data integrity – tool kit & standards			See 4
Demonstration Projects		27. Early child development indicator (EDI) – phase II enable communities to utilise this information	D Hazlehurst /F Oberklaid
		28. Mothers and babies collection-re-invigoration	Alliance
		29. 2006 census modification for retention of records	Alliance /R Madden
		30. Child card data gaps	Alliance
		31. Workplace policies data gaps	Alliance
		32. Weight/Nutrition data gaps	Alliance

Table Five – Agreed Actions on Data Network