Submission Template
2016 National Research Infrastructure Roadmap
Capability Issues Paper

Name	Maureen Turner
Title/role	CEO
Organisation	BioGrid Australia Ltd

Health and Medical Sciences

Question 1: Are there other capability areas that should be considered?

Within the Health and Medical Science capability, the need for real-time data linkage of hospital/clinic based clinical, bio-specimen, pathology, imaging and genomic data (i.e. Real-World Data or RWD) to enable analysis to generate Real-World Evidence (RWE) is not addressed. A critical gap in health data is data on patient outcomes. Whilst outcome registries are very important and valuable, it is expensive and time-consuming to collect these data. Data on patient outcomes exists as part of what is already collected as RWD, but is not easily useable as it is distributed across many healthcare and other organisations.

Real-time data linkage of many health datasets will provide capability that distils RWE on patient outcomes from RWD, and will support a precision medicine approach to health care. It is therefore critical that health and medical research data linkage infrastructure and capability expands beyond the current administrative and registry static data linkage platforms to include real-time linkage of much richer clinical and clinically relevant data to provide evidence on patient outcomes that supports more informed and efficient decision-making in health care. Concomitant with such real-time data linkage infrastructure, capability in analysis of large, linked, dynamic datasets with complex structures will be required to make the best use of RWD. To date, solutions to enable this have not been facilitated, and BioGrid Australia, a fit-for-purpose national research infrastructure capability, can assist.

Support for BioGrid Australia would enable a cost-effective solution to provide access to real-time clinical, biobanking and genomic data that could expand the capabilities of:

- Translating Health Discovery (THD) and Translating Health and Medical Research (THMR) projects via the Australian Therapeutic Pipeline (Pipeline) managed by Therapeutic Innovation Australia (TIA), and
- Data linkage platforms such as the Population Health Research Network (PHRN).

Further investment in infrastructure to support more clinical data linkages would enable Pipeline members to access hospital/clinic-based clinical and genomic data that could enable:

- Pre-clinical innovation through using clinical and genomic data for pre-selection and adding a data validation phase;
- Biobanking through use of clinical and genomic data to enhance the biobank sample with clinical annotation;

- Clinical trials and biomarker biostatistics through use of clinical and genomic data to identify patients for trials, enhance recruitment and provide clinically rich annotated data; and
- Biological and cell therapies platforms through using clinical and genomic data for pre-selection of patients for clinical trials.

With BioGrid Australia a member of the Pipeline, THD projects and Pipeline members would be able to leverage the capabilities of BioGrid Australia to access relevant clinical, biobanking and genomic data to enhance THD projects.

Discussions have commenced with PHRN to identify areas to work collaboratively; thus far, two key areas have been identified and BioGrid Australia will continue to engage with PHRN to expedite expansion of the current administrative and registry static data linkage capability to include real-time linkage of richer clinical and clinically relevant data. With BioGrid Australia and PHRN working collaboratively, data connectedness would allow the whole to be far greater than the sum of the parts, leveraging the individual lines of investment and maximising strengths to benefit Australian researchers through the expansion of national data linkage research infrastructure.

Question 2: Are these governance characteristics appropriate and are there other factors that should be considered for optimal governance for national research infrastructure.

Good governance is critical to the success of any organisation including national research infrastructure. Evidence of appropriate transparency of decision making, management of conflicts of interest, governing body skills, mix, planning, audit and probity should be demonstrated.

Research infrastructure should be governed under a robust governance model that includes the characteristics identified in the issues paper, but Risk Modelling and Management and Innovation Governance should also be considered.

BioGrid Australia has a robust organisational governance structure and operational data governance structure. BioGrid Australia, a not for profit company limited by guarantee with 34 members represented by 66 sites across the nation, is governed by an independent skills based board and receives scientific research direction from a member based advisory committee. BioGrid Australia's data governance model and policies comply with state, territory and commonwealth privacy and use of health data legislation and the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research. BioGrid Australia's data governance is reviewed and agreed to by all members' legal counsel and reviewed and approved by all members' ethics committees before data is linked to the BioGrid Australia platform and made available to researchers. This governance model should be leveraged even further to progress a national approach for clinical data integration that can support Australian medical research.

Question 3: Should national research infrastructure investment assist with access to international facilities?

Yes, it is important to ensure that Australian researchers can access international data and collaborate internationally. A model to support access and utilise international facilities, should

enable affordable access for Australian researchers. It should also increase the impact of Australian research, as national research capabilities can be integrated into international studies.

At present Australia is unable to share health and medical data internationally from a national data infrastructure facility as most research data is fragmented across the country and stored in silos. . National research infrastructure investment should ensure that research infrastructure developed in Australia is compatible internationally. As far as practicable, the data frameworks, standards and interoperability of Australian data should conform to international best practice.

By ensuring that Australian health and medical data are internationally interoperable, Australia can then be nationally interoperable. A commitment at the national research infrastructure level to adopt international standards will have the benefit of filtering down to individual local project level where individual researchers will be more likely to adopt international standards if there is incentive for them to do so. The incentive in this case is access to the well structured national resources.

Question 4: What are the conditions or scenarios where access to international facilities should be prioritised over developing national facilities?

Where there are well-established international collaborations and standards, or well-managed "meta-resources" such as the International Nucleotide Sequence Database Collaboration (INSDC) the UniProt consortium, or The Cancer Genome Atlas (TCGA) Australia should focus on access rather than duplication. Where expensive facilities such as the USA's free electron laser has been developed, attempts should be made to create routes of access, rather than a local facility.

Question 5: Should research workforce skills be considered a research infrastructure issue?

It is well recognised that the health and medical research sector has skills gaps in the quantitative, mathematical and computational sciences including big data analytics, biostatistics and health/bio-informatics. Importantly, a lack of skills in data analytics in the health sector is a barrier to generation of RWE and research that uses large, linked health datasets as described above. Researchers with these skill sets are discouraged from health and medical research as they are uncompetitive for project-based grant funding and there are no funding mechanisms or career paths to support and retain people with these specialist skills. In addition, continuous development of research infrastructure is required to keep it at the cutting-edge internationally. Development and support of career paths in specialist skills is therefore essential for optimal use, development and sustainability of research infrastructure and to maximise the return-on-investment made in research infrastructure.

A research workforce model that encourages multi-disciplinary 'team science' that brings together clinicians, researchers and those with specialist skills in research technology/infrastructure is critical to maximising investment in national research infrastructure. The key element missing is people (not project) funding for researchers with specialist skills and technical specialists.

A multi-disciplinary 'team science' approach is particularly important when research infrastructure is being developed and established. For example, in genomics, specialist resource needs to be skilled to work with genome sequencing systems and pipelines both at an operational

and interpretational level. Without appropriate, and often, very specialised skills, some research infrastructure would not be able to operate and provide services to the research sector. Similarly, clinicians accessing outputs, e.g. variants, from genome sequencing need to be appropriately skilled to understand and interpret these data for clinical application.

Question 6: How can national research infrastructure assist in training and skills development?

National research infrastructure should collaborate with training and education providers particularly in the development of career paths that can support the innovative development and ongoing operations of research infrastructure to maximise use of infrastructure. Major infrastructure investments should be accompanied by co-investment in education and training professionals that can work with technology experts to optimise skills development.

In addition it is critical that national research infrastructure operators develop and provide training materials and information sessions for relevant operational staff and end users. The funding life cycle needs to ensure that training and skills development is addressed, enacted and reviewed at appropriate time intervals taking into account changes in technology, methodology and/or process.

Question 7: What responsibility should research institutions have in supporting the development of infrastructure ready researchers and technical specialists?

Research institutes should be providing relevant training for their researchers, just as any organisation should for their employees. Whilst research institutes themselves may not be able to provide the relevant training, they should be providing the time and budget for relevant researchers to attend training and become appropriately skilled where appropriate, implying that it is not efficient for every researcher to be skilled in everything. For complex research infrastructure subsidised training and access for researchers would assist in ensuring that these complex facilities were used and researchers were skilled as required.

A research workforce model that encourages multi-disciplinary 'team science' that brings together clinicians, researchers and those with specialist skills in research technology/infrastructure is a critical area that research institutions would be able to demonstrate leadership.

In addition there is the opportunity for like institutions to demonstrate leadership by working together to run education/skills training for researchers. Research institutions could commit to 'core service' facilities where a group with relevant expertise can support the broader research community such as the Victorian Life Sciences and Computing Initiative (VLSCI) which has developed deep expertise which is leveraged by many researchers.

Question 8: What principles should be applied for access to national research infrastructure, and are there situations when these should not apply?

A well-structured and transparent model for access to national infrastructure is critical. Opportunities should be sought to engage more with industry where access to important research infrastructure would be of benefit. Co-investment and/or fee for service with industry would provide important funding for the sector. NCRIS collaboration with other agencies that support research activity such as the NHMRC would enable consistency and transparency of a cost model

for access to national research infrastructure. This would also be applicable to other research capability areas. Research grant funding needs to take into account the cost of access to research infrastructure and that national infrastructure may reside in a variety of organisational settings, not just within the academic sector. Nevertheless, it needs to be recognised that research infrastructure will not be able to be fully funded through a cost recovery model without a significant impost on the NHMRC budget, and the NCRIS funding model should support equitable and affordable access for researchers.

Question 9: What should the criteria and funding arrangements for defunding or decommissioning look like?

All initiatives need to budget for start-up costs, ongoing infrastructure renewal, establishing a sustainability model, either partial or full, and if it is relevant or required, defunding or decommissioning activity. Should a wind up of resource and infrastructure be required a clear exit strategy including stakeholder communications and skills management needs to be developed and executed.

It is widely known in the research community that data connectedness is poor and this inhibits efficient and cost-effective research; it is no different in the health and medical sector. The technical solutions that may reside within health institutions to enable data integration that supports clinical decision making often do not support the requirements for research data linkage. Fit-for-purpose research infrastructure that supports data linkage for research use is critical to address the gap in real-time linkage of clinical and clinically-relevant data for research.

The funding of fit-for-purpose research infrastructure should be considered to ensure that the areas identified in the issues paper such as big health data, biobanking and population genomics, national health and medical data capability, and managing and leveraging research data insights can be achieved in a cost-effective manner. Support for BioGrid Australia, a fit-for-purpose national research infrastructure, would enable a cost-effective solution to provide access to real-time clinical, biobanking and genomic data that could expand the capabilities of data linkage platforms such as PHRN.

Question 10: What financing models should the Government consider to support investment in national research infrastructure?

Co-investment with industry or other governments where there is a clear benefit to industry or other countries in our region. Where the national research infrastructure being considered provides value to the research sector, industry or the local region there is a strong case for cooperation.

This is particularly relevant to health and clinical data where industry has on a number of occasions supported infrastructure development to facilitate collection of real world data (RWD) to enable analysis of real world evidence (RWE). National clinical registries in bowel, breast and prostate cancers, sponsored by BioGrid Australia and funded by industry, are used for comparative effectiveness research into patient outcomes and clinician decision making. Once data is collected and provided to industry in aggregated summary form for their purposes, these data remain available through the BioGrid Australia platform for researchers to access.

Through industry investment, data is collected through systems and processes that otherwise would not exist, thus providing RWD through BioGrid Australia that can be analysed for RWE. This is a successful funding model with industry to support the development and maintenance of infrastructure to collect and analyse rich clinical treatment outcome data across the nation. This model could be further leveraged to support collection and analysis of the rich clinical treatment outcome data associated with adverse events, particularly in relation to the use of drugs in the patient population.

In addition, the federal government should consider more favourable tax concessions for philanthropists. This would allow more funding to flow to the research sector and could address the lack of capital funding for research infrastructure in Australia.

Question 11: When should capabilities be expected to address standard and accreditation requirements?

All capabilities should address standard and accreditation requirements if there are such requirements available. This could be a deliverable in the NCRIS funding arrangements for such capabilities. Funding to support preparing for accreditation would be well received and ensure that there were no barriers to accreditation being achieved. Addressing national and international standards and accreditation across Australia's national research infrastructure will ensure there is compliance and consistency across the country that will facilitate international collaboration. In the health and medical science capability, services that provide information that may inform clinical decision making, e.g. genomic pipelines contributing to pathology diagnostic services, should be accredited.

For new and specialised capabilities, part of the support that could be provided by NCRIS (in partnership with the research infrastructure organisation) would be development of sensible and reasonable metrics and standards for measuring the performance of the research infrastructure. This would provide a framework for analysing over time the effectiveness and usefulness of funding provided to research infrastructure organisations.

Question 12: Are there international or global models that represent best practice for national research infrastructure that could be considered?

Question 13: In considering whole of life investment including decommissioning or defunding for national research infrastructure are there examples domestic or international that should be examined?

Question 14: Are there alternative financing options, including international models that the Government could consider to support investment in national research infrastructure?

Health and Medical Sciences

Question 15: Are the identified emerging directions and research infrastructure capabilities for Health and Medical Sciences right? Are there any missing or additional needed?

Within the Health and Medical Science capability, the need for real-time data linkage of hospital/clinic based clinical, bio-specimen, pathology, imaging and genomic data (i.e. Real-World Data or RWD) to enable analysis to generate Real-World Evidence (RWE) is not addressed. RWE will improve patient outcomes and support a precision medicine approach to health care. It is therefore critical that health and medical research data linkage infrastructure and capability expands beyond the current static administrative and registry data linkage platforms to include real-time linkage of much richer clinical and clinically relevant data to provide evidence on patient outcomes that supports more informed and efficient decision-making in health care. BioGrid Australia, a fit-for-purpose national research infrastructure capability, can assist with solutions to this issue.

Support for BioGrid Australia would enable a cost-effective solution to provide access to real-time clinical, biobanking and genomic data that could be leveraged even further to progress a national approach for data integration that can support:

- Translating Health Discovery (THD) and Translating Health and Medical Research (THMR) projects via the Australian Therapeutic Pipeline (Pipeline) managed by Therapeutic Innovation Australia. Data and information are of value to the Pipeline. Of particular importance will be the unique opportunity the Pipeline in Australia offers in yielding RWD about the potential/real impact of new products on real world population groups with multiple morbidities given the current and projected prevalence of chronic complex conditions in an ageing population. In addition, the emerging capabilities in cell biology may well empower new diagnostics capable of driving more personalised medicines. In combination with the established genomics infrastructures, capabilities to interrogate living human tissue stem cells will enhance medical opportunities (e.g. intestinal organoids for cystic fibrosis drug selection).
- Big Health Data, Biobanking and Population Genomics, National Health and Medical Data Capability, and, Managing and Leveraging Research Data Insights. Better data connectedness through support for complimentary research infrastructure capabilities such as BioGrid Australia and PHRN to work collaboratively leveraging the investment to date and maximising strengths to benefit Australian researchers through the expansion of national research infrastructure to enable researchers' better access to individual clinical and population level data.

Some examples of successful research initiatives due to BioGrid Australia's existing federated research infrastructure include:

- Linkage of genomic data to clinical data to better inform diagnosis and treatment. This work has been and continues to be conducted with the Melbourne Genomics Health Alliance, initially through a demonstration project, and now via a Victorian government funded 4-year initiative. This model is being expanded nationally through funding from the NHMRC for the Australian Genomic Health Alliance. Both the Melbourne and

Australian Genomic Alliances collaborate with the Global Alliance in the US, thus interoperability internationally is important.

- Linkage of administrative, registry, clinic and primary care data to create a Cancer Health Data Platform for the Victorian Comprehensive Cancer Centre. This platform is conducting the first large-scale, comprehensive data linkage to link primary care and hospital data for cancer patients in Australia. Through this health services research the VCCC is seeking to improve patient outcomes with evidence-based and cost-effective care.

With appropriate infrastructure support these types of national research initiatives that require real-time linkage of health and medical data would be possible and are critical for the advancement of precision medicine in Australia.

Question 16: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

There are a number of data sharing and data linkage initiatives emerging in the US and Europe. While most of these initiatives have well structured plans and policies for data sharing for research, jurisdictional privacy legislation can prove to be challenging to enable data sharing. Australia should engage with these initiatives to ensure that the issues of international data sharing are addressed within the context of national research infrastructure capability. It is important for Australia to also engage with national health institutes in Europe and the US as well as the Welcome Trust, UK and the Institute of Medicine, US as they further develop their data sharing capabilities.

Question 17: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Health and Medical Sciences capability area?

Underpinning Research Infrastructure

Question 30: Are the identified emerging directions and research infrastructure capabilities for Underpinning Research Infrastructure right? Are there any missing or additional needed?

Question 31: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

Question 32: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Underpinning Research Infrastructure capability area?

Data for Research and Discoverability

Question 33 Are the identified emerging directions and research infrastructure capabilities for Data for Research and Discoverability right? Are there any missing or additional needed?

Question 34: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

Question 35: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Data for Research and Discoverability capability area?

Other comments

If you believe that there are issues not addressed in this Issues Paper or the associated questions, please provide your comments under this heading noting the overall 20 page limit of submissions.

A cost-effective solution to support a national framework for real-time data linkage of hospital/clinic based clinical, bio-specimen, pathology, imaging and genomic data (i.e. Real-World Data or RWD) is required. These capabilities will enable analysis to generate Real-World Evidence (RWE). We know that this is possible due to the success of a well established and proven operational federated national data linkage platform, BioGrid Australia. RWE will improve patient outcomes and support a precision medicine approach to health care. It is critical that health research data linkage infrastructure and capability expands beyond static administrative and registry data linkage to include real-time linkage of hospital/clinic based data to support more informed and efficient decision-making in health care.

Commencing in 2003 and considering other facilities globally, BioGrid Australia is still unique; its data linkage platform with its unique legal and ethical framework should be leveraged even further to progress a national approach for data integration that can support Australian medical research. BioGrid Australia operates a federated infrastructure model for data sharing; the key features of this federated model include:

- BioGrid Australia does not store linked data in a data warehouse: rather the platform is structured to access data across the federation as soon as specific research is authorised.
- Custodians of data linked to the BioGrid Australia platform have full control over whom and for what purpose their data are accessed. This is managed via the BioGrid Australia online Access Request System, which is overseen by the Melbourne Health Human Research Ethics Committee on behalf of the BioGrid Australia collaborators.
- BioGrid Australia utilises either probabilistic identity matching or exact matching using a cryptographic hashing function, to generate a master linkage key, to enable linkage of individuals' records across multiple datasets. This is achieved by using a defined set of personally identifying information to create a BioGrid Australia Unique Subject Identifier (master linkage key) for each individual.
- All data transfers and linkages within BioGrid Australia's infrastructure are performed over secure and encrypted network connections. Identifiable data is received and managed separately to individuals' health information so authorised researchers only access coded data (i.e. with personal identifiers removed).
- BioGrid Australia uses the highly regarded IBM Security Identity Management System to control authorised user provisioning and web access to data available through BioGrid Australia.

BioGrid Australia specialises in real-time linkage of hospital/clinic-based and managed clinical and research data including treatment outcome, genomic, bio-specimen, imaging and patient administration data systems. BioGrid Australia is a trusted and reliable partner with 34 institutions representing 66 sites across Australia. Importantly, BioGrid Australia is not an alternative to other national research infrastructure platforms, rather it facilitates productivity by the provision of capability that is missing, that is, comprehensive real-time ongoing clinical data linkage across the clinical setting in Australia. Most data linkage is project specific. BioGrid Australia is the only collaboration network that provides the federated infrastructure and processes to enable hospital/clinic-based clinical data to be linked and accessed in real-time.

To support its cancer health services research program, the Victorian Comprehensive Cancer Centre (VCCC) is creating a Cancer Health Data Platform. This project will conduct the first large-scale, comprehensive data linkage to link primary care and hospital data for cancer patients. Each of the VCCC partner hospitals has agreed to contribute clinical and administrative data, which will be linked to the Victorian Cancer Registry to enable identification of cancer patients, and then to data from primary care as part of the collaboration with the Health and Biomedical Informatics Centre and the Department of General Practice at the University of Melbourne.

Until now, it has not been possible to extract detailed data for primary care patients, therefore care given to cancer patients by General Practitioners and other primary care health professionals has been invisible to health services research that seeks to improve evidence-based and cost-effective care. This project is now using the BioGrid Australia data sharing infrastructure to facilitate federated data sharing and ethical access to health data for research.

Data and information are of value to researchers and industry. Of particular importance will be the unique opportunity in yielding information about the potential/real impact of new products on real world population groups with multiple morbidities given the current and projected prevalence of chronic complex conditions in an ageing population. Data linkages via BioGrid Australia and other data linkage platforms such as Population Health Research Network (PHRN) and the Australian Institute of Health and Welfare (AIHW) are critical to provide access to individual clinical and population level data. Data connectedness between these capabilities (BioGrid Australia, PHRN, and AIHW) is critical to provide RWD to provide RWE that will facilitate precision medicine.

BioGrid Australia has a contractual agreement with AIHW to provide linkage to the National Death Index for BioGrid Australia collaborators. This partnership enables death data (date and cause of death) to be readily accessed for ethically approved projects to be linked to clinical data available through BioGrid Australia. These data linkage partnerships enhance the accessibility of data for health and medical research thus making research more efficient. Linkages to other important datasets such as Pharmaceutical Benefits Scheme and Medicare Benefits Scheme should be more accessible to researchers therefore leading national research infrastructure should be able to facilitate better connectedness of these data to other clinical, biobanking and genomic datasets.

BioGrid Australia's unparalleled experience in federated real-time data linkage is very highly regarded by the recently funded Australian Genomic Health Alliance (AGHA). BioGrid Australia's work with the Melbourne Genomic Health Alliance linking clinical phenotypic data with genomic data informed what is feasible regarding making relevant disparate data available in the clinic to assist clinician decision making. BioGrid Australia is now a national platform partner of the AGHA.