Adding geospatial data to the data mix

Ming Li, Muktar Ahmed, and David Roder

Cancer Epidemiology and Population Health, University of South Australia

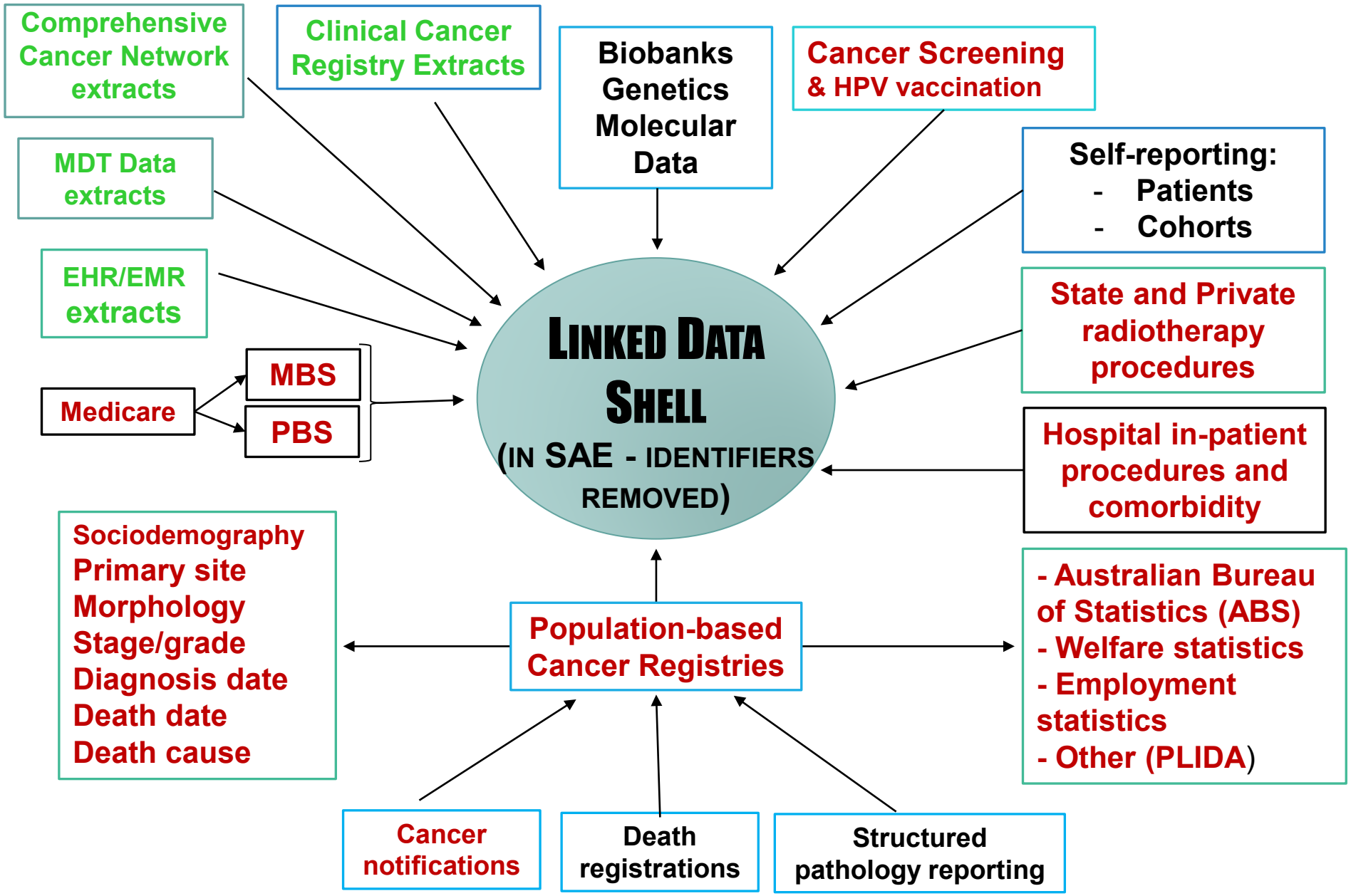
University of South Australia

BREASTSCREEN AUSTRALIA National Conference
Canberra, ACT

March 14, 2024

Geospatial data source: Li M, van Gaans D, Ahmed M, Nguyen AM, Reintals M, Holmes A, Roder D. Determinants of breast screening participation using small-area data in South Australia: gaining past and future insights from geospatial evidence.

(S2) A CANCER DATA FRAMEWORK — POPULATION AND HEALTH SYSTEM COMPONENT



(S3) POTENTIAL DATA FRAMEWORK — inputs, storage, and outputs

DATA INPUTS	DATA STORAGE	DATA OUTPUTS
POPULATION-WIDE (mostly)	**NAT HEALTH DATA HUB**	**Health behaviour**
PBCRs (S&T and other registries)	Other SAE:	**Cancer outputs**
ACD (multiple S&T cancer reg)	DATALAB	Incidence
NDI (death data)	SURE	Mortality
PLIDA (sociodemographic data)	ERICA	Survival
Screening reg. breast, cervix, CRC)	NDII	Prevalence
PROS/PREMS (still in early phase)		Stage & non-stage PIs
Hospital inpatient/ED		Time to treatment
NIHSI		Screening pathways
Radiotherapy		Care pathways
MBS/PBS		NCCIs
Vaccination (AIR)		PROS/PREMS
		Other outputs
		Multimorbidity
		Comorbidity
		Side effects (ST&LT)
SUB-POPULATION REACH (mostly)	**NAT HEALTH DATA HUB (?)**	**Drill-down data detail**
National Health Surveys		
Clinical registry extracts		
ACCN extracts		
MDT extracts		
EMRs		
Genetic/molecular data		
SUPPLEMENTARY DATA	**NAT HEALTH DATA HUB (?)**	**Drill-down data detail**
45&Up		
Melbourne 2020		
Qld Breast Outcomes Study, etc.		

(S4) **Age-standardized biennial participation in BREASTSCREEN: Australian females aged 50-69 years1996-2000 reference set at 100**

1996-2000	54.9% (100)
2000-2005	50.9% (93)
2005-2010	56.2% (102)
2010-2015	54.3% (99)
2015-2019	54.3% (99)

Data source: Cancer Australia NCCI & AIHW/BS Aust reports

(S5) Age-standardized biennial participation in BREASTSCREEN: Australian females aged 50-74 years

Characteristic	2018 & 2019 (A)	2020&2021 (B)	Ratio B/A (% change)
Age (years)			
50-54	**49.8**	**42.8**	**0.859 (-14.1%)**
55-59	52,5	45.3	0.863 (-13.7%)
60-64	57.3	49.4	0.862 (-13.8%)
65-69	59.6	51.9	0.871 (-12.9%)
70-74	56.3	50.0	0.888 (-11.2%)
AUST	54.2	47.0	0.867 (-13.3%)
State & Territory			
NSW	53.8	43.2	**0.803 (-19.7%)**
VIC	53,2	45.8	0.861 (-13.9%)
QLD	54.4	51.4	**0.945 (-5.5%)**
WA	54,4	48.2	0.886 (-11.4%)
SA	58,4	50.3	0.861 (-13.9%)
TAS	59.8	56.0	0.936 (-6.4%)
ACT	58.1	51.9	0.893 (-10.7%)
NT	**38.0**	**34.6**	0.911 (-8.9%)
AUST	54.2	47.0	0.867 (-13.3%)
RESID-REMOTENESS			
Major city	53.1	45.0	**0.847 (-15.3%)**
Inner regional	57.0	50.6	**0.888 (-11.2%)**
Outer regional	56.9	53.7	0.944 (-5.6%)
Remote	53.0	48.8	0.921 (-7.9%)
Very remote	**41.6**	**37.1**	**0.892 (-10.8%)**
AUST	54.2	47.0	0.867 (-13.3%)
SEIFA IRSD			
1 (most disadvantage)	**51.8**	44.8	0.865 (-13.5)
2	55.5	48.3	0.870 (-13.0)
3	53.9	47.2	0.876 (-12.4)
4	55.1	48.2	0.875 (-12.5)
5 (least disadvantage)	54.8	45.9	**0.838 (-16.2)**
AUST	54.2	47.0	0.867 (-13.3)
MAIN LANGUAGE (at home)			
English	56.2	49.3	0.877 (-12.3%)
Other	**45.5**	**37.0**	**0.813 (-18.7%)**
AUST	54.2	47.0	0.867 (-13.3%)
INDIGENOUS STATUS			
Indigenous	**38.3**	**34.9**	**0.911 (-8.9%)**
Non-Indigenous	54.4	47.1	0.866 (-13.4%)
AUST	54.2	47.0	0.867 (-13.3%)

Data source: BREASTSCREEN Australia monitoring reports (AIHW)

(S6) Age-standardized interval cancer rates (0-24 months) per 10,000 women: BREASTSCREEN Australia*

Diagnostic period	First screening round - rate (95% CI)	Subsequent screening rounds - rate (95% CI)
1996-1998	9.5 (8.8, 10.2)	10.3 (9.8, 10.8)
1998-2000	10.7 (9.8 11.6)	10.7 (10.3, 11.2)
2000-2002	9.5 (8.6, 10.6)	10.1 (9.7, 10.5)
2002-2004	8.8 (7.8, 9.9)	9.4 (9.0, 9.7)
2004-2006	9.1 (8.0, 10.3)	9.3 (9.0, 9.7)
2006-2008	**8.7 (7.7, 9.9)**	**9.0 (8.7, 9.4)**
2008-2010	**7.8 (6.8, 8.9)**	**9.0 (8.7, 9.4)**
2010-2012	**8.0 (7.0, 9.1)**	**9.3 (9.0, 9.6)**
2012-2014	**8.1 (6.9, 9.3)**	**9.1 (8.8, 9.5)**
2013-2015	**7.9 (6.9, 9.1)**	**8.9 (8.6, 9.2)**
2016-2018	**8.1 (7.2, 9.1)(est.)**	**9.1 (8.8, 9.4)(est.)**

* Ages: 50-69 years at diagnosis

Data source: BREASTSCREEN Australia monitoring reports (AIHW) Note: later data relate to ages 50-74 years

(S7) Title: Exploratory study of biennial BREASTSCREEN SA participation at SA2-level – key Geospatial predictors

Purpose:
To test the geospatial operational processes and assess the face-value validity of results.

***Outcome variables*:**

SA2-level biennial BREASTSCREEN participation rates (numerators = BREASTSCREEN participants (2014-2015); denominators=census data (2016); for women age 50-74 years). note: years chosen for census proximity.

Candidate predictor variables (examples):

Birth country, mother's education level, employment status, provision of unpaid care for disability, mortgage stress, residential remoteness, relative socioeconomic disadvantage, and other socio-demographic variables. All predictors at SA2-level.

Analysis:

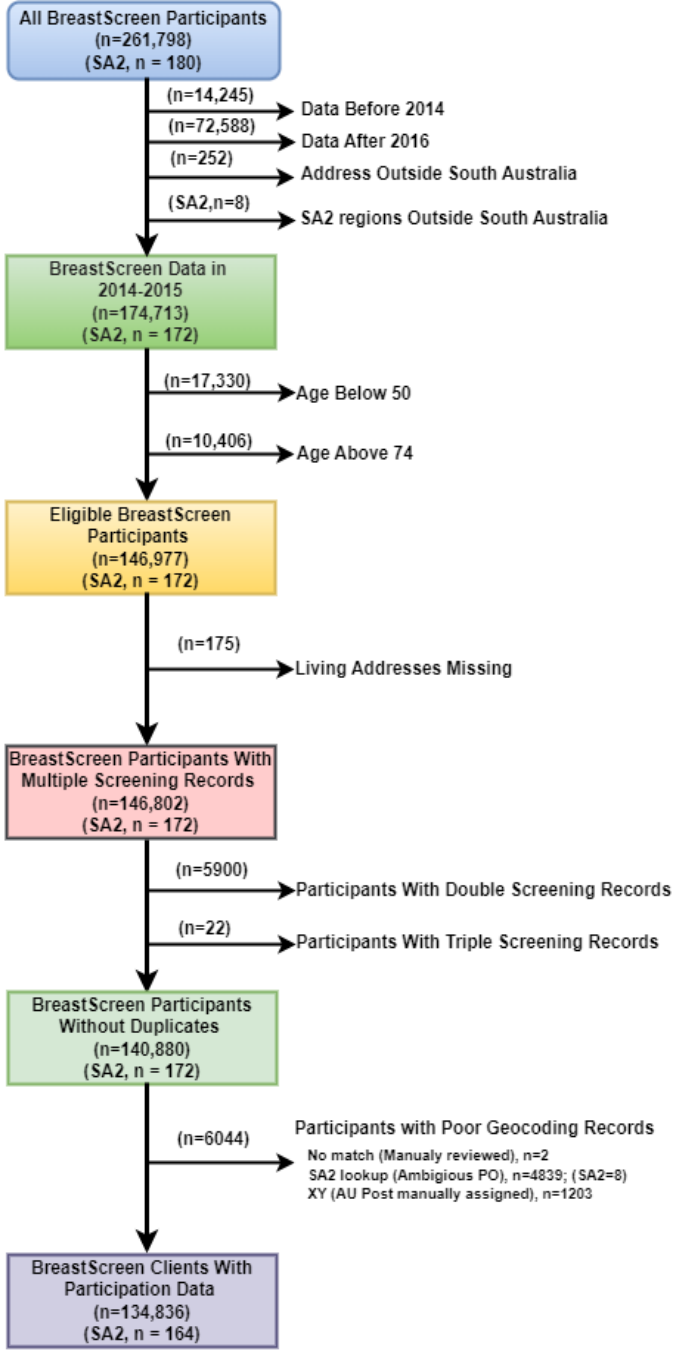
Estimated odds ratios for SA2 participation being below the SA mean derived from multivariable logistic regression with backwards elimination of indicators where p>0.05. Likelihood ratio testing used with STATA 17.0 (STATACORP 2021).

Potential advantages and disadvantages of SA2 level analyses?

Advantages	Disadvantages
Convenience?	Ecological fallacy if generalise to individuals?
Aggregated data more accessible?	Relationships less precisely expressed?
Data often already coded and validated?	Less opportunity to adjust for confounding?
Results probably reflect SA2 combined person and geographic environmental effects	Less opportunity to investigate interactions and effect modification?
Data custodians less defensive?	Reduced statistical power?
Privacy protection greater?	Sub-group analyses not possible?
SA2-level data facilitate targeting of interventions by area?	Courser analyses with reduced opportunity for partitioning?
Can check data quality at SA2-level?	Data quality not assessable at person level?

Data source: Li M, van Gaans D, Ahmed M, Nguyen AM, Reintals M, Holmes A, Roder D. Determinants of breast screening participation using small-area data in South Australia: gaining past and future insights from geospatial evidence. Submission Draft.

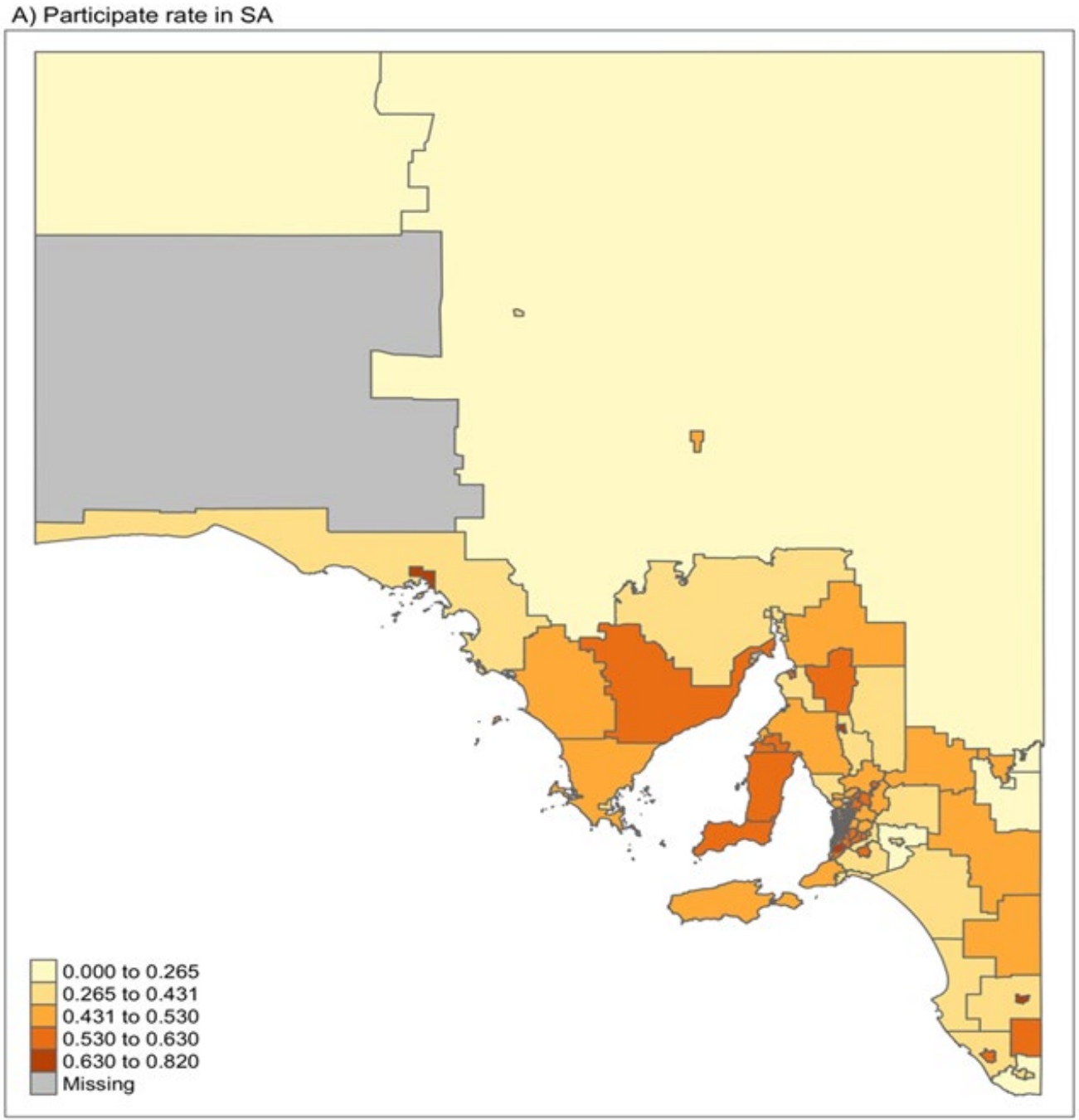
(S8) Geospatial analysis of SA2-level indicators of biennial BREASTSCREEN participation in SA



Data source: Li M, van Gaans D, Ahmed M, Nguyen AM, Reintals M, Holmes A, Roder D. Determinants of breast screening participation using small-area data in South Australia: gaining past and future insights from geospatial evidence. Submission Draft.

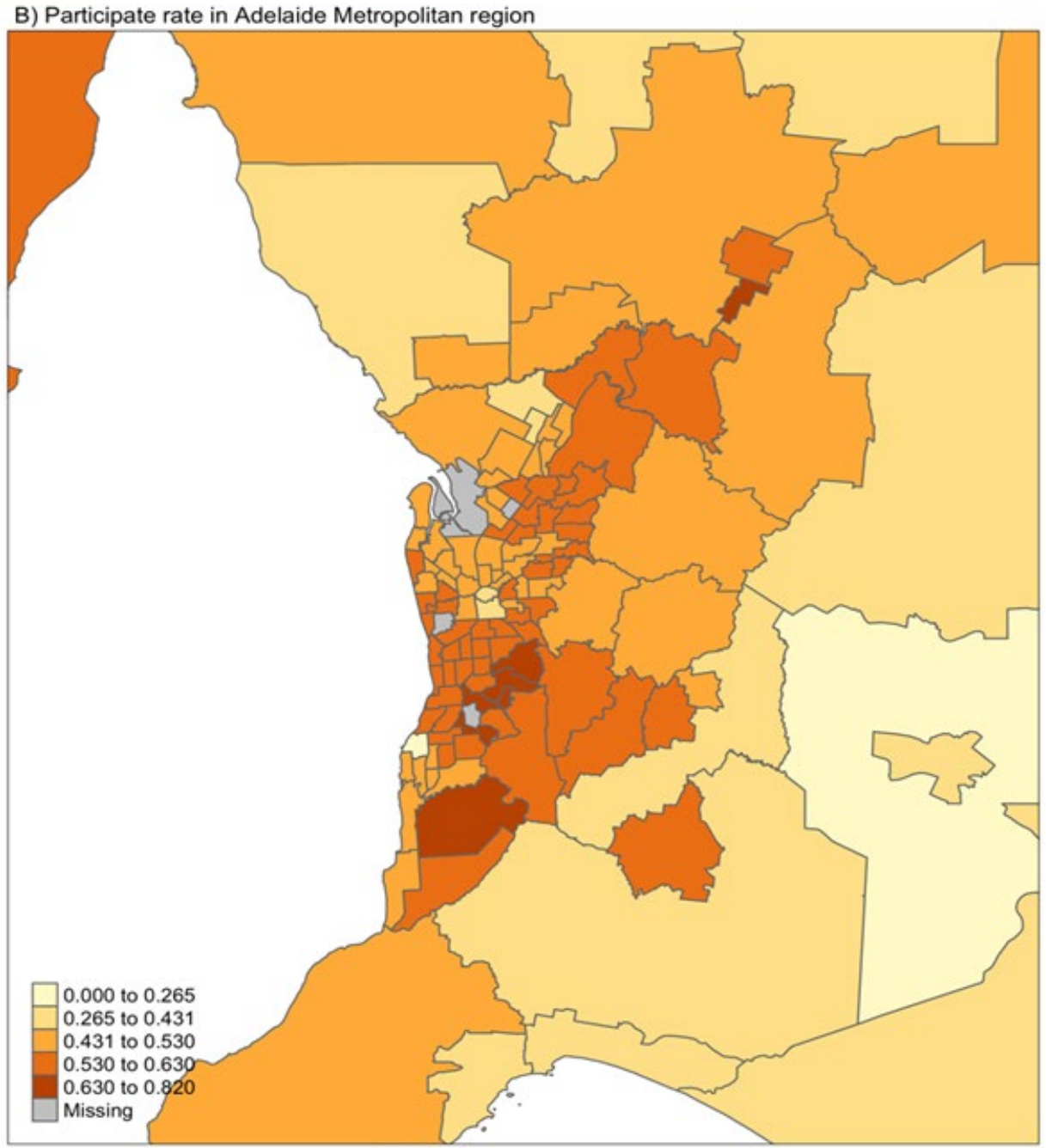
Note: The total number of Statistical Area Level 2 (SA2) regions is sourced from the 2016 Australian Bureau of Statistics (ABS-2016).

(S9) Geospatial analysis of SA2-level indicators of biennial BREASTSCREEN participation in SA



Data source: Li M, van Gaans D, Ahmed M, Nguyen AM, Reintals M, Holmes A, Roder D. Determinants of breast screening participation using small-area data in South Australia: gaining past and future insights from geospatial evidence. Submission Draft.

(S10) Geospatial analysis of SA2-level indicators of biennial BREASTSCREEN participation in metropolitan Adelaide



Data source: Li M, van Gaans D, Ahmed M, Nguyen AM, Reintals M, Holmes A, Roder D. Determinants of breast screening participation using small-area data in South Australia: gaining past and future insights from geospatial evidence. Submission Draft.

(S11) Odds ratios (95% confidence limits) for **low biennial participation** in BREASTSCREEN by key SA2 predictors in 2014-2015: SA women aged 50-74 years*

SA2 Indicator (ordinal thirds)	Unadjusted	Adjusted
Country of birth (% Aust. born)		
Lowest (<70.8%)		**1.00**
Mid (70.8-78.3%)		**0.52 (0.30, 0.88)**
Highest (>78.3%)		**0.27 (0.11, 0.67)**
SEIFA IRSD		
SES Highest		**1.00**
SES Mid		**6.08 (3.45, 10.70)**
SES Lowest		**17.00 (9.84, 29,86)**
Residence remoteness		
Metropolitan		**1.00**
Non-metropolitan		**4.94 (2.30, 10.60)**
Mother education		
Education Highest 1/3	1.00	**Excluded as p>0.05**
Education Mid 1/3	2.77 (1.21, 6.38)	
Education Lowest 1/3	7.59 (2.62, 21.97)	
Employment rate		
Lowest 1/3	1.00	**Excluded as p>0.05**
Mid 1/3	0.48 (0.30, 0.77)	
Highest 1/3	0.19 (0.12, 0.30)	
Provided unpaid disability care		
Lowest 1/3		**1.00**
Mid 1/3		**0.41 (0.24, 0.70)**
Highest 1/3		**0.81 (0.47, 1.39)**
Mortgage/rent stress (in low $ households)		
Lowest 1/3		**1.00**
Mid 1/3		**2.92 (1.68, 5.07)**
Highest 1/3		**6.59 (3.34, 13.00)**

Data source: Li M, van Gaans D, Ahmed M, Nguyen AM, Reintals M, Holmes A, Roder D. Determinants of breast screening participation using small-area data in South Australia: gaining past and future insights from geospatial evidence. Submission Draft.

*Multivariable logistic regression

(S12) Extending SA2 predictors using PLIDA (MADIP) sources (self/parents)?

- **To model:**
  - Cancer screening, vaccination; Cancer incidence, mortality, prevalence, survival; Cancer stage, non-stage prognostic indicators; Time to treatment; Treatment first round, subsequent treatment; Multimorbidity, comorbidity; Side effects; Recurrence markers; patient-reported experiences/outcomes
- **Characteristics – key sociodemographic indices derived from:**

Indigenous status

Aboriginal, Torres Strait Islander, Aboriginal/Torres Strait Islander

CALD Subgroup status

Country of birth; Year of arrival in Australia; years lived in Australia; Ancestry; First language, preferred language; main language spoken at home; Migrant status; Proficiency in spoken English, foreign citizenship

Educational status

Highest educational attainment (school/post school); Educational/training institution type; Vocational education and training; Apprenticeship; Non-school qualification; Student status

Employment

Employment status; Occupation status

Income & support

Personal income; Household/family income; Household/family size; Equivalized household income; Use of social services

Residential relative socioeconomic status (IRSD)

Residential remoteness

Major city; Inner/outer region; Remote/very remote

Health status

Long-term health conditions; Medical disability; Mental condition/disability; Disability status

NOTE: WITH NUMERATOR/DENOMINATOR ALIGNMENT

Data platform: Reference: ABS. Person level integrated Data Asset (PLIDA). Canberra, 2023.

Data sources: ABS; ATO; Depts. Education, Social Services, Home affairs; Services Australia

(S13) Extending SA2 predictors through surveys (e.g., NSW 45 & UP; Melbourne Health 2020; Qld Breast Cancer Outcome Study, 2010-2013)

EXAMPLE: FROM QLD BREAST CANCER OUTCOME STUDY

DIAGNOSTIC AGE & PROGNOSTIC FACTORS

Age at diagnosis, TNM stage at diagnosis, Tumour grade, Triple negative

CLINICAL AND Self EXAMINATION

Clinical breast examination, Breast self-examination

LIFESTYLE

Healthy weight /overweight, Physical activity, Smoking, Drinking

REPRODUCTIVE HISTORY

Age at menarche, Duration of menstruation, Menopause, Age at starting contraceptive use, Duration of contraceptive use, Age at first childbirth/Number of children, Duration of breast feeding, Duration of HRT

FAMILY HISTORY

Relatives with breast or ovarian cancer

INDIVIDUAL SES

Education, Employment, Income, Number of cars, Marital status, Language spoken at home, Private insurance

RESIDENTIAL FACTORS

Residential remoteness, Treatment accessibility, Residential area disadvantage

SCREENING FACILITIES

Screening facility type, Screening facility remoteness, Screening area disadvantage

DATA SOURCE: Kou K et al. Severity and risk factors of interval breast cancer in Queensland, Australia: a population-based study. Breast Cancer 2023; 30:466-477

(S14) # CANCER AUSTRALIA – NATIONAL CANCER CONTROL INDICATOR FRAMEWORK

NCCI FRAMEWORK

INDICATOR CATEGORY (n)	INDICATOR ITEM
Prevention (n=7)	Smoking prevalence
	Alcohol consumption
	Overweight & obesity
	HPV vaccination
	Hepatitis B vaccination
	Sun Exposure
	Diet
	Physical activity
Screening (n=6)	Breast screening rates
	Abnormal breast screen to assessment rates
	Cervical screening rates
	Colorectal screening rates
	Colonoscopy follow-up rates
	Time-bowel screen to diagnostic assessment
Diagnosis (n=2)	Cancer incidence
	Stage at diagnosis
Treatment (n=4)	Surgery
	Radiotherapy
	Systemic therapy
	MDT supported care
Psychosocial care (n=2)	Patient experience
	Screening for distress
Research (n=1)	Clinical trial participation
Outcomes (n=9)	Cancer mortality
	Relative survival
	Conditional relative survival
	Survival by stage (childhood)
	Mortality to incidence ratio
	Survival by stage (all ages)
	Prevalence
	Recurrence
	Palliative care

Data source: https://ncci.canceraustralia.gov.au

(S15)

COMMENTS/ PROPOSALS

1. A National Linked Data Resource should be developed in manageable steps to support implementation of the Australian cancer plan cancer, service delivery and research
2. This Resource would be designed to ensure more timely and efficient availability of data for service monitoring and research at person and geospatial level (e.g., SA2 level)
3. Data access should be available to analysts at national and S&T level, and in academia, for approved projects
4. The linked data with identifiers removed should be stored/analysed in a National Data Hub or alternative SAE
5. The data Governance Structure should reflect funder, service provider, and consumer interests, and ensure Indigenous Data Sovereignty
6. A formalised process should enable data users to include supplementary data for approved projects