Cause For Analysis: Using A Causal Analytic Workflow For Data Processing In Antimicrobial Resistance Surveillance

Michelle L Harrison[1,2*^], Mengyao Tang*[3], Alkim Ozaygen[1], Paul Turner[4,5], Thomas Snelling[1,2], Phoebe CM Williams[1,2,6], Yue Wu[1,2]

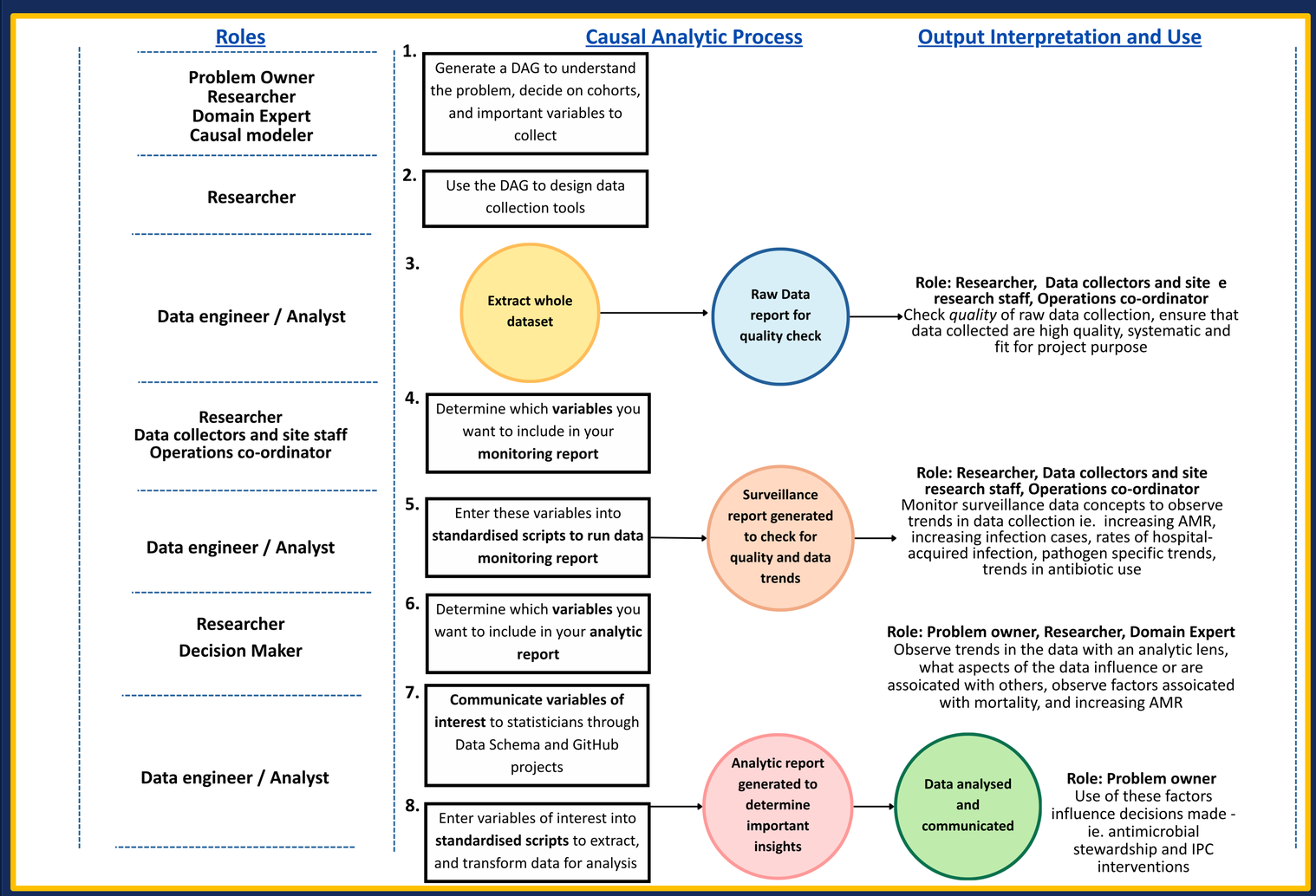
[1]*Sydney School of Public Health, Faculty of Medicine, The University of Sydney,* Sydney, Australia [2]*Sydney Infectious Diseases Institute*, Sydney, Australia [3]*Columbia University, School of Public Health, NY, New York* [4]*Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford,* UK [5]*Cambodia Oxford Medical Research Unit, Angkor Hospital for Children,* Cambodia [6]*UNSW School of Women and Children's Health, UNSW,* Sydney, Australia *Co-first authors

BACKGROUND

Efforts have been made to **standardise the collection of antimicrobial resistance (AMR) data globally.[1-4]** There remain challenges in improving data-collection efficiency, in processing to facilitate more accurate comparison of high-quality AMR data between institutions and across countries,[5] and in translating analytic insights into better decision-making. Neonates and children are significantly under-represented in these data and overcoming processing challenges will be essential.

Challenges encountered in the processing of AMR and infection surveillance data include:

- Differing raw data structures
- Complexity in the scientific and medical domains that lead to differences in identification and management of likely contaminant
- Managing antimicrobial prescription and reporting antimicrobial susceptibility testing (AST) results, which collectively lead to a high time burden to process data, and inefficiency in communication across the project team.

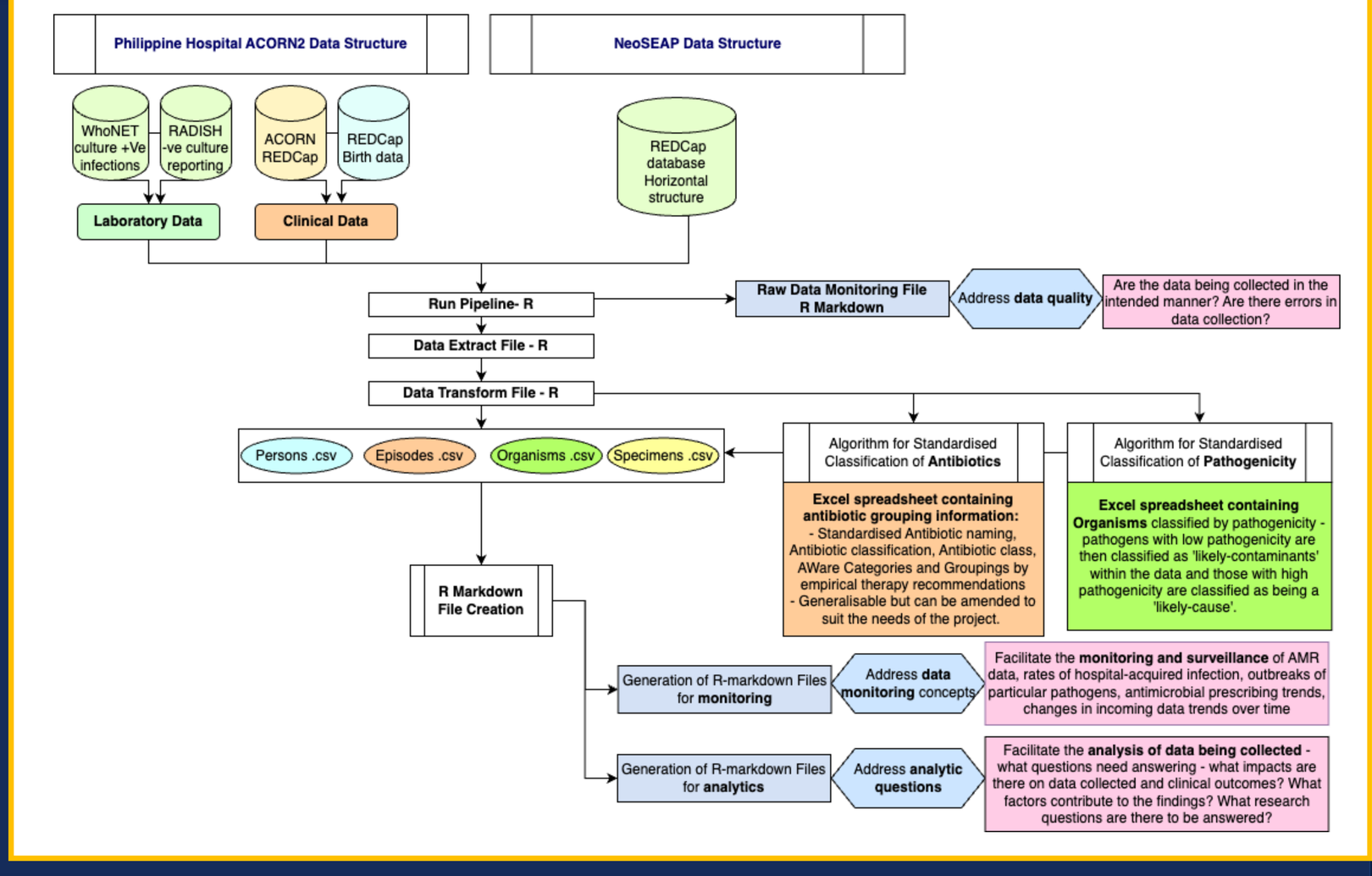


METHODS

- We aimed to design an approach to standardise data processing and provide a structure for efficient coordination across disciplines to achieve quality data monitoring and analytics using a generic research data structure.
- We developed an 8-step causal analytic process that integrates multidisciplinary input from data analysts, decision makers, domain experts, data collectors, and researchers to streamline data processing.
- We used two different neonatal infection surveillance databases (ACORN2 and NeoSEAP) to illustrate how researchers can apply the causal analytic process to achieve transparent and efficient analytics.

RESULTS

- We implemented standardised scripts to convert AMR data from differing structures (wide and long) into a generic relational data structure representing four main concepts: 'persons', 'episodes', 'organisms', and 'specimens'.
- We designed and implemented standardised algorithms to group and interpret antimicrobial and AST data, and to categorise pathogens by pathogenicity.
- This process led to the production of three data monitoring reports – a raw data report for data quality and validity checks, a monitoring report for identifying key patterns in the data, and an analytic report for answering important research questions.



CONCLUSION

More efficient methods are essential for processing AMR data. This causal analytic process will provide improved structure for reliable, transparent, generalisable and comparative reporting of AMR across clinical sites globally.

Acknowledgments: Authors would like to acknowledge the work done by the clinical data collection sites for the NeoSEAP prospective study and the ACORN2 study at the Philippine General Hospital. We would like to acknowledge funding support from the National Health and Medical Research Council (NHMRC) of Australia.

References