Project Details	
Project Code	MRC22IIARBa Sheppard
Title	Attributing the source of antimicrobial resistant diarrheal pathogens in African children
Research Theme	Infection, Immunity, Antimicrobial Resistance & Repair
Summary	Diarrhoeal disease is a major cause of mortality among children in low-income countries. Joining a large MRC funded program you will collect and sequence metagenome samples to quantify the relative contribution of different antimicrobial resistant pathogens to human infection. Time spent in Bath, Bristol and The Gambia will help understand transmission networks, and bioinformatics and machine learning risk models will identify effective interventions.
Description	SIGNIFICANCE: The WHO ranks diarrhoeal disease as the second most common cause of mortality among children under five years of age in low and middle-income countries (LMICs), accounting for 10.6 million annual deaths in this age group. In this program, we will address the paucity of systematic diarrhoeal studies for countries that have the highest child mortality rates. Building on MRC funding and an extensive collaborative network, we are now in the position to understand the epidemiology and source of multiple enteric pathogens in LMICs and apply this knowledge to help those most at risk. RESEARCH WITH IMPACT: Factors such as household crowding, poor sanitation, contaminated water and cohabitation with animals, may constitute transmission risks, but their relative importance is unknown. This is a major concern, particularly as frequent or chronic enteric (re)infection is linked to significant morbidity and growth faltering in children. Comparative genomics offers a solution for untangling complex transmission networks. Gut bacteria adapt to their hosts and these adaptations are reflected in their genome. Therefore, by sequencing the genome of strains from humans and infection reservoirs it is possible to identify genomic source markers and determine where clinical cases originated. INTERDISCIPLINARY TRAINING THROUGH COLLABORATION: This project will deliver excellent interdisciplinary training with time spent at three partner institutions including the MRC Unit The Gambia (clinical microbiology), Bath (microbial genomics/bioinformatics) and Bristol (Machine learning). The student will join a large well-funded (MRC, BBSRC) international collaborative network and attend meetings, workshops and conferences. PROJECT WORKPLAN: 1. Sampling (The Gambia): Using established protocols, local technical leads will collect sample swabs from symptomatic humans (through hospitals) and the environment (e.g., potable water, food, domestic and wild animals, etc). Campylobacter, Salmonella, Shigella, Klebsiella, E. coli and Vibrio will be identified using laboratory microbiology techniques and antimicrobial resistance (AMR) status will be determined. 2. Genomics and metagenomics: Culture-dependent single strain sequencing and metagenomics approaches will be used to investigate the co-occurrence of enteropathogen species and strains. For metagenomics, current methodologies in the Sheppard Laboratory will identify multiple species directly from primary samples (eg. faeces) and from plate sweeps. Genomes and metadata will be archived in project-specific BIGSdb databases (https://sheppardlab.com/resources), containing >100,000

	enteric pathogen genomes for comparison. 3. Identifying source-specific genomic markers: We have developed quantitative genome-wide association studies for enteric bacteria. Using these evolutionary models, the student will identify genes (including AMR) that are overrepresented among samples from different countries, animals, food production systems, and humans (clinical/asymptomatic). 4. Quantitative source attribution: The genetic elements (genes, alleles, SNPs) that are more frequent in genomes from certain sources constitute markers for determining the likely origin of infecting strains. Using machine learning and probabilistic source attribution models, isolates will be quantitatively assigned to putative source based on comparison of genome-wide segregating markers, allowing transmission networks to be determined. Subsequent risk calculations can describe reservoir source-sink dynamics for specific genes (including AMR). 5. Evidence-based interventions (Bristol): Novel artificial intelligence and statistical data science methodologies will be used to link metadata to assess the impact of potential interventions. Using the relative importance of different sources, and the cost of intervention, monetized decision support tools will identify most effective intervention points in the transmission network.
Supervisory Team	
Lead Supervisor	
Name	Professor Samuel Sheppard
Affiliation	Bath
College/Faculty	Science
Department/School	Biology and Biochemistry
Email Address	s.k.sheppard@bath.ac.uk
Co-Supervisor 1	
Name	Professor Andrew Dowsey
Affiliation	Bristol
College/Faculty	Bristol Veterinary School
Department/School	Department of Population Health
Co-Supervisor 2	
Name	Dr Ben Pascoe
Affiliation	Bath
College/Faculty	Science
Department/School	Biology and Biochemistry
Co-Supervisor 3	
Name	Dr Jahangir Hossain
Affiliation	Other
College/Faculty	Science
Department/School	Disease Control and Elimination
Co-Supervisor 4	
Name	
Affiliation	
College/Faculty	
Department/School	