Abstract

Background: The relationship between asymptomatic *Salmonella* exposure within the gastrointestinal tract and *Salmonella* bacteraemia is poorly understood, in part due to the low sensitivity of stool culture, and the lack of validated molecular diagnostic tests for the detection of *Salmonella* in stool. The study aimed to determine a reliable molecular diagnostic test for *Salmonella* in stool specimens.

Methods: We optimized an in-house monoplex real time polymerase chain reaction (PCR) for the detection of *Salmonella* TTR and InvA genes in stool by including a selenite broth pre-culture step for *Salmonella* before DNA extraction, and validated their specificity against other local common pathogens. Then we assessed their performance against a well-validated multiplex PCR targeting the same TTR and InvA genes, and against stool culture using clinical stool specimens collected from a cohort of 50 asymptomatic healthy Malawian children that were sampled at 1-month intervals over a period of 12 months. We employed a latent Markov model to estimate the specificities and sensitivities of PCR methods.

Results: TTR and InvA primers were both able to detect all the different *Salmonella* serovars tested, and had superior limits of detection if DNA was extracted after selenite pre-culture. TTR sensitivity and specificity for monoplex-PCR were (99.53%, 95.46%)
and for multiplex-PCR (90.30%, 99.30%) respectively. InvA specificity and specificity for using monoplex-PCR was (95.06%, 90.31%) and multiplex-PCRs (89.41%, 98.00%) respectively. Sensitivity and specificity for standard stool culture were 62.88% and 99.99% respectively. Culture showed the highest PPV (99.73%) and mono-TTR had the highest NPV (99.67%).

**Conclusion:** Test methods demonstrated high concordance although stool culture and monoplexed TTR primers had superior specificity and sensitivity respectively. The use of selenite pre-enrichment step increased *Salmonella* detection rate. Taken together, molecular detection methods used here could be used to reveal the true extent of both asymptomatic and symptomatic *Salmonella* exposure events.

**Keywords**
Salmonella Typhi, nontyphoidal Salmonella, bacteremia, gastrointestinal tract, diagnostics, stool culture, polymerase chain reaction

---

This article is included in the Malawi-Liverpool Wellcome Trust Clinical Research Programme gateway.
Introduction

Salmonellae cause a huge global burden of morbidity and mortality. They are globally estimated to be responsible for 300,000 deaths and Salmonella enterica serovars Typhi and Paratyphi A are the predominant cause of invasive Salmonella infections in south and southeast Asia, and cause between 129,000 to 223,000 global deaths per year. In contrast, non-typhoidal Salmonella (NTS) serovars, principally S. Typhimurium and S. Enteritidis, are a common cause of invasive disease in sub-Saharan Africa (sSA) and in blood. In 2017, NTS caused an estimated 535,000 cases with sSA having the highest incidence. Risk factors for invasive NTS (iNTS) disease include young age, recent malaria, and advanced HIV disease. Case fatality rates for iNTS in young children, people infected with HIV, and living in the sSA region were estimated at 13.5%, 41.8% and 15.8% respectively. This is in marked contrast to the presentation of Salmonella disease in high-income countries, where NTS typically cause a self-limiting diarrhoeal disease in healthy individuals, while bloodstream or focal infections are rare and mainly occur in individuals with specific risk factors such as diabetes, neoplastic and autoimmune disease, or immunosuppressive therapy. It is notable, however, in both settings, that invasive NTS disease in adults and children are not always associated with diarrhoea.

We previously described in under-five year-old children the sequential development of cellular and humoral immunity against the Salmonella serovars causing iNTS disease, and that acquisition of this immunity is associated with decreasing incidence of disease, suggesting that this immunity is protective. Previous studies have reported that healthy young children experience transient asymptomatic episodes of gastrointestinal infection with non-typhoidal Salmonella, and we therefore hypothesise that episodes of asymptomatic Salmonella exposure in the healthy gastrointestinal tract during early childhood may facilitate the development of protective immunity. Balanced against this beneficial effect of exposure, diarrhoeal disease results from enteric Salmonella exposure, and invasive NTS disease also follows episodes of asymptomatic gastrointestinal exposure in susceptible children including those with malaria or malnutrition, or immunocompromised individuals.

Elucidating the relationship between Salmonella exposure events within the gastrointestinal tract and resultant Salmonella immunity or Salmonella disease is thus critical for understanding the pathogenesis of iNTS disease. Lack of affordable and rapid diagnostic tools for the detection of bloodstream and intestinal Salmonella disease hampers our understanding of Salmonella disease epidemiology and pathogenesis. Blood culture is considered the gold standard diagnostic test for Salmonella bacteremia and is highly specific but has a number of drawbacks; poor turn-around time of between 2 to 7 days, and low sensitivity of about 20% - 30% for samples collected 7 days post-infection. Molecular detection of Salmonella in blood also has limited apparent sensitivity, and different assays are in development.

Stool culture is similarly considered the gold standard test for the detection of Salmonella in the intestinal tract. However, stool culture, even for diarrhoeal disease when the bacterial load is likely to be high, has poor sensitivity (<50%), and is labour and time consuming. Real-time PCR has a short turnaround time and is potentially highly sensitive compared to standard culture, and has the capacity for automation and testing for multiple targets. However, stool PCR test performance is hindered by PCR inhibitors and a large number of genetically closely related enteric bacteria. These pose a challenge in the generation of highly specific and sensitive primers for real-time PCR (qPCR) for Salmonella. Furthermore, a lower infective load of Salmonella colonization during asymptomatic infection may further limit detection by PCR.

With this background, we validated an in-house monoplex qPCR method for the detection of Salmonella in stool specimens, and compared them with a validated multiplex based qPCR and standard stool culture. Both qPCR assays used primers and probes based on the Salmonella tetrahydronate respiration gene (ttr), and the Salmonella invasion gene A (invA). Stool specimens were collected from healthy, mainly asymptomatic healthy Malawian children aged 6 – 18 months. Assessing the performance of a diagnostic test is challenging when the existing "gold standard" test being used has known low sensitivity or specificity. Statistical methods, such as the Latent Markov model, are used to assess the performance of diagnostic tests without assigning a gold standard test. Since the current "gold standard" is known to lack sensitivity, we employed a latent Markov model, in order to estimate the specificities and sensitivities of PCR methods without assigning a gold standard.

Methods

Description of study participants and specimens

Stool specimens collected from a longitudinal cohort of children aged 6 – 18 months who were recruited from Zingwawwa Health Centre (ZHC) in Blantyre, Malawi, were used to compare the performance of molecular and standard culture for detection of Salmonella in stool. The main study started recruitment in August 2013 and follow up was concluded in December 2014. Group sensitization of the study, by well-trained study nurses, was done to parents or guardians of six month old children attending a vaccination clinic at ZHC. Individual sensitization was also done to parents or guardian that were interested in joining the study. Children who met the inclusion criteria of being healthy were recruited into the study after obtaining consent. Children born preterm (less than 38 weeks' gestation), HIV positive or HIV exposed, and those with fever >38°C or any acute illness were excluded from the study.

Stool samples were collected monthly until they were aged 18 months. Stool specimens were collected in sterile and clean containers and transported to the laboratory on the same day. From 60 children who were recruited at 6 months of age, 10 children withdrew from the study, and 600 stool specimens were collected and tested by culture, on the day of sample.
collection at the College of Medicine and Malawi Liverpool Wellcome Laboratory. Molecular tests were done on frozen samples that were available at the time the tests were done.

Salmonella stool culture
A matchstick head-size sample of stool was inoculated in selenite F broth (Oxoid, UK, catalog number: 2300631) and aerobically incubated overnight at 37 °C for 18–24 hours. The top layer (1 ml) of an overnight culture was spun at 20,000 g for 5 minutes and the pellet was sub-cultured on Xylose Lysine Deoxycholate (XLD) agar (Oxoid, UK, catalog number: 2547703). An aliquot of the selenite broth was also frozen for molecular detection (below). Presumptive Salmonella colonies were cultured onto sheep blood agar (Oxoid, UK, catalog number: 2910831) and MacConkey agar plates (Oxoid, UK, catalog number: 2529552) and incubated aerobically at 37°C for 18–24 hours. Salmonella colonies were then distinguished from other enteric bacteria (i.e. Citrobacter and Serratia) using triple sugar iron agar (Oxoid, UK, catalog number: 1882283) and Urea agar (Oxoid, UK, catalog number: 1779617) biochemical tests. Further Salmonella identification was determined using API® 10S (bioMérieux, France, catalog number: 1007181060) according to the manufacturer’s instructions.

Monoplex-qPCR TTR and InvA assay
Validation of the monoplex- qPCR TTR and InvA assay.
For the monoplex-qPCR, the TTR primers and probe were designed and validated by Federal Institute for Risk Assessment, Berlin, Germany, according to the published DNA sequence of the S. enterica serotype Typhimurium ttr locus for Salmonella detection (GenBank accession no. AF282268) which had demonstrated high specificity and sensitivity rates when used on food samples21. The InvA DNA primers and probe were previously designed in-house in Malawi for blood samples17. Both primers required optimisation for use in stool specimens. The DNA sequence of all the primers and probes used in this study are listed in Table 1.

Specificity of TTR and InvA primer/probe set for Salmonella compared to other local pathogens
To determine the specificity of the primers for Salmonella compared to other local pathogens, 9 different locally isolated and whole genome sequenced Salmonella strains and 26 pure isolates of non-Salmonella bacterial strains locally isolated from blood culture were tested using TTR and InvA primer/probe sets (Table 2). These strains were chosen because they are genetically closely related to Salmonella or because their growing conditions are similar to Salmonella. These strains were collected from MLW bacterial blood culture repository. Overnight cultures of the frozen samples were made on SBA or LB agar. One colony was then cultured in liquid media. After reaching stationary growth phase, a known and matched concentration of about 10⁶ CFU was used for DNA extraction using QIAamp Fast DNA Stool Mini Kit (QIAGEN, Netherlands, catalog number: 51604) but without the bead beating step. Miles and Misra technique was used for bacteria quantification.

Limits of detection in different conditions
A well-characterized invasive S. Typhimurium ST313 strain (D23580), isolated from an HIV negative child in Malawi, and representative of our commonest invasive blood stream infections, was used as a reference strain for determining limits of detection in varying kinds of sample22,23. Three types of Salmonella sample were prepared for comparison using RT-PCR: 1) pure Salmonella

| Primer name | Primer direction | Primer code/Probe description |
|-------------|------------------|-------------------------------|
| 1 InvA      | Forward          | AGCGTACTGGAGGGAAGGGAAAG      |
| 2 InvA      | Reverse          | CACCGAAATACCGCAATAAAG        |
| 3 InvA      | Probe            | Fam-TCACGGTCTTCTTGACGGTGCGAT |
| 4 TTR       | Forward          | CTCACCCGGAGATTTACACAATGG     |
| 5 TTR       | Reverse          | AGTCGACACCAAGGATGACCACATC   |
| 6 TTR       | Probe            | 6FAM-CACCGACGGGAGACGCGACTTT  |
| 7 InvA-TAC  | Forward          | GGCAATTCGTTATCGGCGGATA      |
| 8 InvA-TAC  | Reverse          | CACGGTGCAATTAGAGAAGACAAACA   |
| 9 InvA-TAC  | Probe            | CCGCCCGGATGCGTGGGTT         |
| 10 TTR-TAC  | Forward          | TCCACCGAGAGATACACCATGG      |
| 11 TTR-TAC  | Reverse          | AGCTCACAGCAAAGTGACCACATC    |
| 12 TTR-TAC  | Probe            | CACCGACGGCAGACCGACTTT      |

Table 1. List of primers and probes sequences used in this study. Primers and probes sequences used in this study include inhouse designed InvA, TTR previously validated for Salmonella detection in food, and TAC-InvA and TAC-TTR used on a well validated TAC assay as pan Salmonella primers.
isolates picked from a blood agar plate, 2) Salmonella cultured in selenite broth and 3) Salmonella spiked into stool. Salmonella stool spiking in stool was done to determine the inhibitory effect that stool may have on the assay which could then affect the limit of detection. For this, a stool sample was collected from healthy individual and confirmed Salmonella negative by culture. The stool sample was thereafter diluted with PBS (50% w/v) and then spiked with S. Typhimurium, D23580 at varying doses of viable bacteria. The viable dose of Salmonella was adjusted across a range from to $10^0 – 10^6$ CFU/ml, and quantified using Miles and Misra technique. DNA was extracted for RT-PCR, as above. All experiments were repeated three times on different days, by the same operator.

Detection of Salmonella in clinical samples using monoplex-qPCR TTR and InvA assay. The primer/probe sets were then used to detect Salmonella in clinical stool samples collected from the longitudinal cohort study of healthy asymptomatic children. For the monoplex qPCR, approximately 200μl top layer of frozen Selenite F broth overnight stool culture, or 200 mg of stool was suspended in 500 μl of PBS and DNA was extracted using QIAamp Fast DNA Stool Mini Kit (QIAGEN, Netherlands, catalog number: 51604) according to the manufacturer’s instructions, with an added bead-beating step. Eluted DNA was stored at –20°C.

A previously-optimised in-house PCR protocol was used \textsuperscript{19}. Briefly, the master-mix for RT-PCR was prepared using pre-defined quantities. A total of 20μl master-mix for each sample was comprised of the following: 12.5μl Platinum® Quantitative PCR Super Mix-UDG (Life Technologies, USA, Catalog number: 11730025), 0.10μl specific forward primer, 0.10 specific reverse primer, 0.10 specific probe (all primers and probes at 200nM), 0.05μl ROX reference dye (Life Technologies, USA, Catalog number: 12223012) at 50nM final concentration, and 7.15μl nuclease-free water. This mixture was transferred to

---

Table 2. Bacterial organisms tested for the specificity of TTR and InvA primer/probe sets. Bacterial organisms used in this study to test for the specificity of TTR and InvA primer/probe sets. Nine Salmonella and 26 non-Salmonella isolates that were previously isolated at MLW laboratory were retrieved tested either as direct or selenite sub-cultured isolates.

| Bacteria isolates                  | Number tested | Direct TTR Positive | Direct InvA Positive | Selenite sub-cultural TTR Positive | Selenite sub-cultural InvA Positive |
|-----------------------------------|---------------|---------------------|----------------------|------------------------------------|------------------------------------|
| Morganella morgana                | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Streptococcus pneumonia           | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Staphylococcus aureus             | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Citrobacter                       | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Klebsiella                        | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Enterobacter                      | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Acinetobacter                     | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Enterobacter intermedius          | 1             | 0                   | 0                    | 0                                  | 0                                  |
| Enterococcus fecium               | 1             | 0                   | 0                    | 0                                  | 0                                  |
| E. coli                           | 17            | 0                   | 0                    | 0                                  | 0                                  |
| S. Typhi                          |               | 1                   | 1                    | 1                                  | 1                                  |
| S. Typhimurium                    | 1             | 1                   | 1                    | 1                                  | 1                                  |
| S. Enteritidis                    | 1             | 1                   | 1                    | 1                                  | 1                                  |
| S. Braenderup                     | 1             | 1                   | 1                    | 1                                  | 1                                  |
| S. Virchow                        | 1             | 1                   | 1                    | 1                                  | 1                                  |
| S. Bonn/Fann                      | 1             | 1                   | 1                    | 1                                  | 1                                  |
| S. Oesterbro/Zanzibar             | 1             | 1                   | 1                    | 1                                  | 1                                  |
| S. Heidelberg                     | 1             | 1                   | 1                    | 1                                  | 1                                  |
| S. Dublin                         | 1             | 1                   | 1                    | 1                                  | 1                                  |
96-well plate PCR wells. 5μl of test DNA, positive controls DNA (DNA from D23580), technical extraction negative control and assay negative control (UV treated water) were added in triplicates to appropriate wells containing 20μl of master-mix. The qPCR was run for 40 cycles using Applied Biosystems® 7500 Real-Time PCR Systems (Life Technologies, USA). The following cycling conditions were used; initial denaturation at 95°C for 1 minute, denaturation at 95°C for 15 seconds, annealing/extension at 60°C for 30 seconds, final extension: 12°C. The threshold was set in the lag phase. An assay was considered to have passed when the positive controls were positive and both the technical extraction negative and assay negative controls were negative. Test sample cycle threshold (Ct) values were evaluated after subtracting the baseline value. Samples with cycle threshold (Ct) values of less than or equal to 35 were considered positive.

Detection of Salmonella using multiplex qPCR assay
As a comparator, we used a well-validated TAC assay on DNA extracted from stool samples, according to the manufacturer’s protocol. The customized Taqman Array Card assay developed and validated at the University of Virginia was used, which has multiple targets for different enteric pathogens, including InvA and TTR targets for Salmonella detection.

To extract total nucleic acid (TNA) from the clinical samples for TAC assay, we used the same DNA extraction kit and protocol that were used to extract whole-stool DNA for the monoplex qPCR assay, with the addition of internal extraction positive controls. For TNA extraction, each sample was extracted together with internal positive controls, Phocine Herpesvirus (PHhv) and MS2. PHhv and MS2 were added to the inhibition buffer before being added to each sample, as previously described. An assay was considered to have passed when both MS2 and PHhv internal positive (amplification crossing the threshold) and negative controls (no amplification crossing the threshold line) passed, and when the sample had sigmoid curve that crossed the threshold line. Only results for Salmonella are reported here.

Statistical analysis
Data were recorded and analysed in MS Excel (version 16.14.1 (18061302)). Sensitivities and specificities of the different PCR methods were estimated using a latent Markov model (LMM). We have previously described the LMM and various extensions that we considered for modeling longitudinal test data. We implemented the LMM within a Bayesian framework using R (version v3.5.1) and JAGS (version 4.3.0) via the rjags (version 4.6) R package. LMMs have been extensively used for discrete-time longitudinal data in the absence of a gold standard diagnostic procedure. We considered several LMMs, with and without mixed effects and with either time-homogeneous or time-heterogeneous transition matrices. Convergence and identifiability of the LMM were checked by inspecting trace plots and computing Gelman-Rubin potential scale reduction factors. The more complex models exhibited poor mixing or convergence of MCMC chains (most likely due to the sparse number of positive samples) and as a result the LMM we used for this dataset is a basic LMM with no random effects and a time-homogeneous transition matrix. To report positive predictive values (PPV) and negative predictive values (NPV), we calculated an estimate of the infection prevalence. For the Bayesian LMM, we report maximum a posteriori (MAP) parameter estimates together with 95% credible intervals (Crl), specifically the highest posterior density intervals (HDI) with 95% coverage. All other analyses report (frequentist) parameter estimates and corresponding 95% confidence estimates (CI).

Ethical considerations
Ethical approval for this work was granted by the University of Malawi, College of Medicine Research Ethics Committee (P.01/13/1327). Written informed consent was obtained from the parent or guardian of each participating child.

Results
TTR and InvA primers for Salmonella do not cross-react with closely related enteric micro-organisms
We first validated the TTR and InvA primers that were used in the monoplex-qPCR assay, by assessing the sensitivity and specificity of the primers for Salmonella, using a standardized number of 10⁵ – 10⁷ CFU/ml of 9 different locally-relevant Salmonella strains, and 26 non-Salmonella bacterial strains as indicated in Table 2. We included 17 strains of E. coli because of the close genomic relatedness of Salmonella and E. coli. Bacterial isolates that were either enriched in Selenite broth (referred here as selenite sub-cultured) or not (referred here as direct culture) were used in this evaluation. We found that TTR and InvA assays both achieved 100% sensitivity and specificity either as direct isolates or selenite sub-cultured isolates. Table 2 demonstrates that all Salmonella strains tested positive with both monoplexed primer pairs, and all other bacterial strains were negative, confirming a lack of cross-reactivity.

Selenite broth culture enhances the detection of Salmonella in stool using either TTR or InvA primers
The limits of detection (LOD) of qPCR for Salmonella were then determined using S. Typhimurium strain D23580 serially diluted, and tested as direct isolates, selenite broth cultured samples, or isolates spiked into a culture-negative stool specimen. We found that limits of detection for TTR were 1, 10 and 100 CFU/ml, and for InvA were 1, 100 and 100 CFU/ml for selenite sub-cultured broth, direct isolates and stool-spiked isolates respectively, with 98.5% qPCR efficiency for TTR and 97.2% qPCR efficiency for InvA. No statistically significant difference was observed in the LOD when TTR was compared with InvA in either direct isolates (p = 0.3212), selenite sub-cultured samples (P = 0.2534), or salmonella spiked stool samples (P = 0.2361). Importantly, we found that the TTR assay was significantly different when direct isolates (LOD = 10 CFU/ml) were compared with selenite sub-cultured samples (LOD = 1 CFU/ml) (p<0.0001), and when selenite sub-cultured isolates were compared to Salmonella spiked stool (p <0.0001), and there was no significant difference when direct isolates were compared to Salmonella spiked stool (p=0.2965).
Similarly, we found that detection in InvA qPCR assay direct isolates was significantly different compared to selenite broth cultures isolates ($p < 0.0001$), and selenite subculture isolates were also significantly different to *Salmonella* spiked stool ($p < 0.0001$), while no significant difference was found between direct isolates compared to *Salmonella* spiked stool samples ($p = 0.2862$). In summary, we found that selenite broth overnight liquid culture of stool samples enhanced the molecular detection of *Salmonella* using either TTR or InvA primers, even if culture of the broth remained negative.

TTR and InvA primers had both high specificity and sensitivity rates, whilst stool culture had high specificity but low sensitivity.

The samples from healthy children were used to determine the performance of stool culture, monoplex TTR, monoplex InvA, multiplex TAC TTR, and multiplex TAC InvA. Standard stool culture was performed on a total of 600 specimens at different time points. Molecular tests were used to detect *Salmonella* in the available 421 stool DNA specimens. We detected *Salmonella* in 23, 40, 29, 56, and 47 of 421 stool specimens, using standard stool culture, TTR, InvA, TAC-TTR, and TAC-InvA respectively. Of the 23 *Salmonella* stool culture-positive samples, 21 samples were also positive with either one or more molecular tests whilst 2 were negative with molecular tests.

Based on a time-homogeneous LMM without random effects (Table 3 and Figure 1A) we reported the specificities and sensitivities of the detection methods with their 95% credible intervals (Bayesian confidence intervals). The observed specificity rates from highest to lowest were for stool culture (99.99%), TAC-TTR (99.30%), TAC-InvA (98.00%), monoplex TTR (95.46%) and monoplex InvA (90.31%) respectively. The observed sensitivity rates from highest to lowest were monoplex TTR (99.53%), monoplex InvA (95.06%), TAC-TTR (90.30%), TAC-InvA (89.41%) and stool culture (62.88%) respectively (Table 3 and Figure 1A).

|                      | Sensitivity | Specificity | Positive predictive value | Negative predictive value |
|----------------------|-------------|-------------|---------------------------|---------------------------|
|                      | MAP (95% Crl) | MAP (95% Crl) | MAP (95% Crl) | MAP (95% Crl) |
| Stool culture        | 0.6288 (0.3916,0.8223) | 0.9999 (0.9949,10000) | 0.9973 (0.8668,10000) | 0.7238 (0.6135,0.8389) |
| TTR                  | 0.9953 (0.8315,1.0000) | 0.9546 (0.9317,0.9749) | 0.5615 (0.3897,0.7275) | 0.9967 (0.8501,10000) |
| InvA                 | 0.9506 (0.7950,10000) | 0.9031 (0.8702,0.9311) | 0.3521 (0.2233,0.4915) | 0.9536 (0.8147,10000) |
| TAC-TTR              | 0.903 (0.6628,10000) | 0.993 (0.9797,0.9987) | 0.8597 (0.6798,0.9736) | 0.9033 (0.7367,0.9869) |
| TAC-InvA             | 0.8941 (0.6721,0.9869) | 0.98 (0.9618,0.9928) | 0.7228 (0.5079,0.8757) | 0.8807 (0.7459,0.9828) |

**Figure 1.** Maximum a posteriori probability estimates of the specificities and sensitivities (Figure 1A), positive and negative predictive values (Figure 1B) together with 95% highest density credible intervals (segments) and posterior density estimates (contours) for stool culture, TTR, InvA, TAC-TTR and TAC. Big dots and error bars represent the median values and 25 and 75 percentile.
High negative and positive concordance for stool culture, monoplex TTR, monoplex InvA, Multiplex TTR, and multiplex InvA

Next, we explored correlations between stool culture, monoplex TTR, monoplex InvA, Multiplex TTR, and multiplex InvA. In this exploration, we considered all test results, whether positive or negative. To account for both censored observations and the longitudinal nature of the data, we calculated repeated measures of correlation coefficients using the ranks of observations for each test (akin to a repeated-measures Spearman correlation coefficient) for measuring the correlation between the Ct values for the four molecular tests and point biserial correlation coefficients based on ranks for measuring correlations between standard stool culture and each of the qPCR tests (Figure 2A). The correlation coefficients vary quite widely from 0.12 (monoplex InvA and TAC-InvA) to 0.8 (stool culture and TAC-TTR). Given that for truly negative samples, the Ct values are effectively randomly distributed near the threshold used to discriminate between positive and negative samples, and that most samples were negative in most tests, the somewhat weak correlations we observe can be driven by the random Ct values for negative samples. For this reason, using only the binary negative / positive outcomes for each test, we computed positive (Figure 2B) and negative (Figure 2C) concordance: for example, in Figure 2B, the intersection of the row labelled ‘TTR’ and the column labelled ‘InvA’ lists the proportion of positive test results for the TTR test that are also positive for the InvA test. Unexpectedly (given that most samples were negative), negative concordance (Figure 2C) was very high, with the lowest negative concordance being 89%. Results for positive concordance (Figure 2B) are also quite high, though there is more variation, ranging from 25% (for positive InvA results confirmed by positive stool cultures) to 100% (positive stool cultures confirmed by positive monoplex TTR or positive monoplex InvA).

Stool culture had high positive predictive value while molecular tests methods had high negative predictive values

To report PPV and NPV, for an estimate of prevalence, we use the model-estimated stationary (time-homogeneous model) probability of being infected (MAP 5.25%, 95% credible interval [3.27%, 8.14%]). From highest to lowest, the estimated PPVs were culture (99.73%), TAC-TTR (85.97%), TAC-InvA (72.28%), mono-TTR (56.15%), mono-InvA (35.21%). From highest to lowest, the estimated NPVs were mono-TTR (99.67%), mono-InvA (95.36%), TAC-TTR (90.33%), TAC-InvA (88.07%) and culture (72.38%) as indicated in Table 3 and Figure 1B.

Discussion

The burden of asymptomatic gastrointestinal exposure to Salmonella which could be linked to either the development of immunity, or conversely to blood-stream infection is not known, due to lack of robust Salmonella detection methods for stool specimens. This study aimed to optimize detection methods, and to validate and compare the performance of monoplex TTR and InvA qPCR assays (TTR and InvA), against TTR and InvA qPCR assays on a validated multiplex qPCR platform (TAC-TTR and TAC-InvA), and compare all molecular methods to standard Salmonella stool culture. Validation of the monoplex TTR and InvA primers showed that the primers do not cross react with other enteric pathogens, and LOD testing showed that selenite pre-culture promotes molecular detection, even when culture is negative. Stool culture demonstrated the highest specificity but low sensitivity than all the molecular tests. Stool culture, despite having low sensitivity, still...
remains important in Salmonella diagnosis. Culture allows for antimicrobial susceptibility testing and strain typing. TTR detected on the monoplex platform demonstrated superior sensitivity to stool culture, InvA, TAC-TTR, and TAC-InvA. All the test methods however displayed high concordance to each other.

Several studies have developed Salmonella detection methods based on antigen detection or nucleic acid amplification\(^{57,31,32,33}\). Both monoplex and multiplex nucleic acid amplification-based detection methods have been developed\(^ {14-37}\). Most of these have however focused on Salmonella detection in blood as opposed to stool specimens. Some multiplex qPCRs to specifically detect Salmonella and its serovars, or for the detection of multiple enteric pathogens in stool specimen (including Salmonella) have recently been developed\(^ {34-38}\). The advantage of multiplex qPCR is that it is fast in determining the main etiological agent in cases where the outcome is caused by multiple pathogens or different serovars, but it is expensive if one is interested in detecting only one particular pathogen. By contrast, the advantage of a monoplex test is that it is economical. In this study, the same primer/ probe sets were tested using both the monoplex and multiplex qPCR platforms. The monoplex qPCR maximized sensitivity while the multiplex panel provided a balanced pay-off between sensitivity and specificity. The high sensitivities of the monoplex qPCR could be attributed to the use of selenite pre-cultured stool as opposed to extraction of DNA from neat stool samples which is used in the multiplex qPCR. Selenite sub-cultured stool samples were not be used on the multiplex platform because the manufacturer’s protocol was followed. Other studies have, however, also demonstrated superior performance of monoplex qPCR when compared with multiplex qPCR. The monoplex qPCR is therefore ideal for studies that are only interested in determining the presence or absence of Salmonella whilst capitalizing on the sensitivity of the test while multiplex qPCR will have an added advantage if a study wants to detect multiple pathogens whilst having a pay-off between sensitivity and specificity.

The TTR primer/ probe set used in the monoplex qPCR was previously validated for use in food samples and required validation in stool specimens. Our in-house developed InvA primer/ probe set also required validation. Both assays demonstrated that they can detect all the different Salmonella strains including S. Enteritidis, S. Typhimurium, and S. Typhi strains which are the commonly isolated strains in Malawi and sSA\(^ {41}\). Comparing the limits of detection of different Salmonella isolate conditions demonstrated that selenite pre-culture achieves a significantly lower limit of detection (1 CFU/ml) as opposed to direct isolates (10 CFU/ml) and Salmonella-spiked stool (10 CFU/ml). Selenite F broth is a selective broth that suppresses the growth of fecal coliforms and streptococci in order to optimize Salmonella growth\(^ {32}\). The LOD achieved after sub-culturing samples in Selenite enrichment broth agrees with results demonstrated by other studies including a study done by Boer et al. who demonstrated that sub-culturing samples in Selenite F broth promotes the recovery of Salmonella in stool samples and improves sensitivity if samples are subsequently tested using molecular methods like PCR\(^ {32,43}\).

We used an LMM to estimate the specificities and sensitivities of the 5 Salmonella detection methods. Stool culture demonstrated the highest specificity but had the lowest sensitivity. All molecular assays; TAC-TTR, TAC-InvA, TTR, and InvA, demonstrated high specificity and sensitivity rates. Compared to the other methods, the monoplex based qPCR TTR achieved the best sensitivity-specificity trade-off as it demonstrates near-perfect sensitivity (99.53%) and still achieves high specificity (95.43%). All molecular test methods had significantly higher sensitivities than stool culture. High specificity and low sensitivity rate for culture have been widely reported\(^ {18}\). Such low sensitivity rates should be taken into consideration when evaluating diagnostic tests. It is clear that a reference test with poor sensitivity is not adequate to evaluate alternative test methods. In such a situation alternative means of evaluating the assays should be used such as the LMM that has been used here. LMMs, and their counterpart for cross-sectional data, latent class models (LCMs), have been used to evaluate diagnostic tests for different pathogens including Salmonella\(^ {44}\).

PPV and NPV vary depending on the prevalence of the condition being tested in any particular population. Our samples were collected from a population that was considered healthy and asymptomatic at the time of recruitment. Using the model-estimated stationary probability of being infected, we estimated the Salmonella infection prevalence of 5.25% in this population. With this prevalence estimate, stool culture demonstrated a high PPV when compared to molecular tests that had high NPVs. When prevalence is low, a small change in specificity will have significant effects on the PPV. Higher PPVs could be observed in a situation where prevalence is high such as when using a cohort of hospitalized diarrheal cases, or during a diarrheal outbreak.

Molecular methods had higher sensitivity but lower specificity, relative to stool culture. The loss in specificity is small compared to the gain in sensitivity and, in the case of Salmonella, the public health cost of false-negative results could be higher if the infection becomes potentially life-threatening due to withholding or delay of treatment. With the high sensitivity, molecular methods were able to detect asymptomatic Salmonella events, critical for the research questions we hoped to pose in this cohort. All the events that were detected here were asymptomatic in healthy children, which are potentially very important in transmission or the development of immunity. The detection of low bacterial burden events could also be relevant in settings like Malawi where unprescribed over-the-counter antibiotic procurement and use is common. Studies that have reported on risk factors of having a culture-negative result has indicated that antibiotic usage before sample collection is the main risk factor. Using molecular techniques such as PCR could overcome this challenge because it detects bacterial DNA regardless of the viability of the pathogen. This
might increase the probability of identifying the infection and reduce sample processing time which could then with proper patient management and treatment if needed.

Our study has several limitations. One main limitation is the use of different sample types for the two qPCR platforms. The use of selenite sub-cultured stool samples in monoplex qPCR may have contributed to the superior performance when compared with the multiplex qPCR. We used neat stool samples for multiplex qPCR to comply with the manufacturer’s protocol. Other studies have however demonstrated that testing primer/ probe sets in the monoplex platform perform better than in the multiplex qPCR platform. Clinical samples used to test the performance of the test are a limitation especially in determining the PPV and NPV. Clinical samples used in the study were collected from a cohort of children that were asymptomatic to Salmonella and remained healthy for most of the one-year study period. Using samples from participants with a clinical diagnosis of Salmonella or diarrhea would improve the PPV and NPV.

Conclusion
The data presented here demonstrate that the addition of selenite pre-enrichment step increases Salmonella detection in stool samples, and that TTR and InvA primer and probe sets used are able to detect different Salmonella strains. The ability of TTR to detect Salmonella with such high levels of specificity and sensitivity when tested using clinical samples collected from a cohort that was mostly healthy, make it a promising assay that could be used for research surveillance studies. The assays could be very useful in studying the transmission of Salmonella infections. This method may perform with different sensitivity and specificity in a chronic carriage, diarrhoeal or invasive Salmonella disease state, since the load and culturability of the pathogen within the stool may be different, and further validation studies would be needed.

We established that selenite pre-culture increased diagnostic yield for molecular detection and identified TTR primers as molecular tools that could best help to reveal the true extent of Salmonella exposure events within the gastrointestinal tract. This will allow us to understand their importance to diarrhoeal and invasive disease pathogenesis and epidemiology in the future.

Data availability
Underlying data
Figshare: Data and software code for Bayesian mixed latent Markov models for binary diagnostic data, https://doi.org/10.6084/m9.figshare.12911870.v3.
1. gitMarcH-Bayesian-mixed-latent-Markov-models-for-binary-diagnostic-data.zip (software code for Latent Markov Model used in this study)
2. Data files used by the uploaded software code:
   - salexpoLIMSDataSetComplete.csv (Date of sample collection and follow-up visit number)
   - TACResults_4Mar TAC TTR TAC InvA Ct For Correlation. csv (Ct values for TAC_TTR and TAC_InvA)
   - TTR & InvA master file Ct for correlation.csv (Ct values for monoplex TTR and InvA)
   - TTRInvASensitivity20170724_corrected.csv (Combined binary results for stool culture, TTR, InvA, TAC_TTR and TAC_InvA used to calculate sensitivity, specificity and correction of the test methods)
3. Raw data:
   - TACResults_TAC-TTR_TAC-InvA_I_Ct_ValuesTAC Results_TAC-TTR_TAC-InvA_IC_Ct-values.csv (raw Taqman array card assay results for test and control sample)
   - Salmonella_Detection_Stool_TTR_InvA_raw_data.xlsx (raw data for the monoplex qPCR assay. Includes results for test and control sample)

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgments
The authors are very thankful to Sr. Rose Nkhata who helped with recruitment and following up of the study participants.

References

1. Crump JA: Progress in Typhoid Fever Epidemiology. Clin Infect Dis. 2019; 68(Suppl 1): S4-s9.
2. Ao TT, Feasey NA, Gordon MA, et al.: Global burden of invasive nontyphoidal Salmonella disease, 2010(1). Emerg Infect Dis. 2015; 21(6): 941-949.
3. GBD 2017 Typhoid and Paratyphoid Collaborators: The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis. 2019; 19(4): 369-381.
4. GBD 2017 Non-Typhoidal Salmonella Invasive Disease Collaborators: The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis. 2019; 19(12): 1312-1324.
5. Crump JA, Lubzy SP, Mintz ED: The global burden of typhoid fever. Bull World
salmonella in Malawi. JSM Proceedings Vancouver Canada. 2018.

Reference Source

27. Plummer M: rjags: Bayesian Graphical Models using MCMC. R package version 4.6-2016. Reference Source

28. Koukounari A, Donnelly CA, Moustaki I, et al.: A latent Markov modelling approach to the evaluation of circulating cathodic antigen strips for schistosomiasis diagnosis pre- and post-praziquantel treatment in Uganda. PLoS Comput Biol. 2013; 9(12): e1003402.

Published Abstract | Publisher Full Text | Free Full Text

29. Koukounari A, Moustaki I, Grassly NC, et al.: Using a nonparametric multilevel latent Markov model to evaluate diagnostics for trachoma. Am J Epidemiol. 2013; 177(9): 913-922.

Published Abstract | Publisher Full Text | Free Full Text

30. Gelman SA, Rubin DB, Rubin DB: Inference from Iterative Simulation Using Multiple Sequences. Stat Sci. 1992; 7: 457-472.

Publisher Full Text

31. Bakdach J, Marusch LR: Repeated Measures Correlation. Front Psychol. 2017; 8: 456.

Published Abstract | Publisher Full Text | Free Full Text

32. Kuipers LM, Chung P, Peeters M, et al.: Diagnostic accuracy of antigen-based immunochromatographic rapid diagnostic tests for the detection of Salmonella in blood culture broth. PLoS One. 2018; 13(3): e0194024.

Published Abstract | Publisher Full Text | Free Full Text

33. Bugarel M, den Bakker H, Grout J, et al.: CRISPR-based assay for the molecular identification of highly prevalent Salmonella serotypes. Food Microbiol. 2018; 71: 8-16.

Published Abstract | Publisher Full Text | Free Full Text

34. Pouzo S, Tamnay AM, Pouzo DA, et al.: Clinical Evaluation of a Multiplex PCR for the Detection of Salmonella enterica Serovars Typhi and Paratyphi A from Blood Specimens in a High-Endemic Setting. Am J Trop Med Hyg. 2019; 101(3): 513-520.

Published Abstract | Publisher Full Text | Free Full Text

35. Heymans R, Vila A, van Heerwaarden CAM, et al.: Rapid detection and differentiation of Salmonella species, Salmonella Typhi, and Salmonella Enteritidis by multiplex quantitative PCR. PLoS One. 2018; 13(10): e0206316.

Published Abstract | Publisher Full Text | Free Full Text

36. Scopes E, Screen J, Evans K, et al.: Evaluation of the Thermo Scientific RapidFinder™ Salmonella Species, Typhimurium and Enteritidis Multiplex PCR Kit. JAOAC Int. 2018; 101(4): 1095-1100.

Published Abstract | Publisher Full Text | Free Full Text

37. Park SH, Ricke SC: Development of multiplex PCR assay for simultaneous detection of Salmonella genus, Salmonella subspecies I, Salm. Enteritidis, Salm. Heidelberg and Salm. Typhimurium. J Appl Microbiol. 2015; 118(1): 152-60.

Published Abstract | Publisher Full Text | Free Full Text

38. Hamnet I, Engbro AL, Parejo J, et al.: Multicenter evaluation of the new QIStat Gastrointestinal Panel for the rapid syndromic testing of acute gastroenteritis. Eur J Clin Microbiol Infect Dis. 2019; 38(11): 2103-2112.

Published Abstract | Publisher Full Text

39. Martin A, Perez-Ayala A, Chaves F, et al.: Evaluation of the multiplex PCR Allplex GI assay in the detection of bacterial pathogens in diarrheal stool samples. J Microbiol Methods. 2018; 144: 33-36.

Published Abstract | Publisher Full Text

40. Beckman AK, Ferreira P: Prospective Investigation of an Automated PCR/Nucleic Acid Microarray-Based Platform for Enteric Pathogen Testing. Lab Med. 2019; 50(4): 390-395.

Published Abstract | Publisher Full Text | Free Full Text

41. Gilchrist JJ, MacLennan CA: Invasive Nontyphoidal Salmonella Disease in Africa. EcoSal Plus. 2019; 8(2).

Published Abstract | Publisher Full Text

42. Boer MD, de Boer RF, Lameijer A, et al.: Selenite enrichment broth to improve the sensitivity in molecular diagnostics of Salmonella. J Microbiol Methods. 2018; 157: 59-64.

Published Abstract | Publisher Full Text | Free Full Text

43. Gomone LW, Mazzulli A, Trimi X, et al.: Evaluating the preservation and isolation of stool pathogens using the COPAN FecalSwab™ Transport System and Walk-Away Specimen Processor. Diagn Microbiol Infect Dis. 2019; 94(1): 15-21.

Published Abstract | Publisher Full Text | Free Full Text

44. Moore CE, Pan-Ngum W, Wijedoru LPM, et al.: Evaluation of the diagnostic accuracy of a typhoid IgM flow assay for the diagnosis of typhoid fever in Cambodian children using a Bayesian latent class model assuming an imperfect gold standard. Am J Trop Med Hyg. 2014; 90(1): 114-120.

Published Abstract | Publisher Full Text | Free Full Text

45. Henrion MYR: Data and software code for Bayesian mixed latent Markov models for binary diagnostic data. figshare. Dataset. 2020. http://www doi.org/10.6084/m9.figshare.12911870.v3

Published Abstract | Publisher Full Text | Free Full Text
Open Peer Review

Current Peer Review Status: ??

Review Version 1

Reviewer Report 27 January 2021

https://doi.org/10.21956/wellcomeopenres.17920.r41999

© 2021 Li Q. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Qiuchun Li

Key Laboratory of Prevention and Control of Biological Hazard Factors (Animal Origin) for Agri-food Safety and Quality, Ministry of Agriculture of China, Yangzhou University, Yangzhou, China

This study used four different methods (monoplex- qPCR TTR, monoplex- qPCR invA, multiplex TAC-TTR, multiplex TAC InvA) in detecting Salmonella in stool samples with comparison to the traditional stool culture methods. The manuscript is not well-written and hard to follow. I have some major questions:

1. Application of these methods into 421 stool samples, 40, 29, 56, and 47 samples were detected positive using mono-TTR, mono-InvA, TAC-TTR, and TAC-InvA, respectively. While 23 samples were detected positive using the stool culture method. the study concluded the detection methods using TTR primers could be used for research surveillance studies, but didn't confirm which method is reliable and applicable? And why?

2. When using the mon-TTR and mono-InvA, the genomic DNA used is from the selenite broth cultured samples, which is time-consuming for cultivation of the stool samples, so what is the advantage of these methods?

3. The specificity of the molecular methods used only 9 Salmonella serotypes, why not use some frequently reported human nontyphoidal Salmonella serotypes like S. London, S. Rissen, and S. Meleagrindis, S. senftenberg, etc.? In addition, it is better to give the background of the prevalence of TTR and InvA in Salmonella serotypes based on genome sequencing analysis. The TAC is also used multiplex PCR methods, there is no detailed information on TAC in the manuscript.

4. The PPV and NPV have also been used to evaluate the molecular detection methods, under the condition of an estimated 5.25% of Salmonella in asymptomatic people, which method is better based on analysis of PPV and NPV?

5. It is necessary to revise the English grammar of the manuscript.
Is the rationale for developing the new method (or application) clearly explained?
Partly

Is the description of the method technically sound?
Partly

Are sufficient details provided to allow replication of the method development and its use by others?
Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Prevalence of Salmonella in Foods and humans.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 21 January 2021

https://doi.org/10.21956/wellcomeopenres.17920.r41992

© 2021 Zhang J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

JY Zhang
National Institute for Communicable Disease Control and Prevention, Beijing, China

The authors optimized an in-house monoplex real-time PCR for the detection of *Salmonella* *ttr* and *invA* genes in stool, validated their specificity, and assessed their performance using clinical stool specimens collected from a cohort of 50 asymptomatic healthy children. The main problems:

1. Nucleic acid detection of enrichment culture can obtain the highest sensitivity, and the sensitivity of culture method is poor, which many studies have mentioned. The authors need to summarize the innovation points again to make it more attractive.

2. Unlike other studies, this study used continuously collected specimens from cohort. If the authors can analyze the queue data, it will be more meaningful.
3. When calculating the sensitivity and specificity of a method, it is necessary to determine whether the tested object is true positive or true negative. Whether the specimen is really positive can be determined by test. The author's conclusion was based on statistical analysis rather than experimental results, which I don't think is ideal.

Minor:
- **Introduction**: Gene names should be in italics and lowercase.
- **Methods**:
  - *Salmonella stool culture*: The volume of selenite F broth should be given. Did this method follow a literature or is it used in the authors' laboratory? Inoculating the centrifuged pellet onto XLD, I think there are too many bacteria, which may affect the subsequent selection of suspected colonies.
  - *Table 1*: The fluorescent groups and quenching groups of the probes needed to be given.
  - *Statistical analysis*: I don't understand the method used by the authors, so I don't comment on it. I suggest that the authors should explain the reliability of this method.
- **Results and discussion**: The analysis of cohort data is more meaningful. The reasons for the inconsistent results should be analyzed.

**Is the rationale for developing the new method (or application) clearly explained?**
Yes

**Is the description of the method technically sound?**
Yes

**Are sufficient details provided to allow replication of the method development and its use by others?**
Partly

**If any results are presented, are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions about the method and its performance adequately supported by the findings presented in the article?**
Partly

**Competing Interests**: No competing interests were disclosed.

**Reviewer Expertise**: Molecular detection of intestinal pathogens. I'm not familiar with the
statistical method mentioned by the author, so I suggest other reviewers review it again.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.