Sanitary inspection of wells using risk-of-contamination scoring indicates a high predictive ability for bacterial faecal pollution in the peri-urban tropical lowlands of Dar es Salaam, Tanzania

Douglas Mushi, Denis Byamukama, Alexander K.T. Kirschner, Robert L. Mach, K. Brunner and Andreas H. Farnleitner

ABSTRACT

Sanitary inspection of wells was performed according to World Health Organization (WHO) procedures using risk-of-contamination (ROC) scoring in the peri-urban tropical lowlands of Dar es Salaam, Tanzania. The ROC was assessed for its capacity to predict bacterial faecal pollution in the investigated well water. The analysis was based on a selection of wells representing environments with low to high presumptive faecal pollution risk and a multi-parametric data set of bacterial indicators, generating a comprehensive picture of the level and characteristics of faecal pollution (such as vegetative *Escherichia coli* cells, *Clostridium perfringens* spores and human-associated sorbitol fermenting Bifidobacteria). ROC scoring demonstrated a remarkable ability to predict bacterial faecal pollution levels in the investigated well water (e.g. 87% of *E. coli* concentration variations were predicted by ROC scoring). Physicochemical characteristics of the wells were not reflected by the ROC scores. Our results indicate that ROC scoring is a useful tool for supporting health-related well water management in urban and suburban areas of tropical, developing countries. The outcome of this study is discussed in the context of previously published results, and future directions are suggested.

Key words | microbial source tracking, peri-urban area, risk of contamination scoring, sanitary inspection, standard and alternative faecal indicator bacteria, well water

Douglas Mushi
Department of Biological Sciences,
Sokoine University, P.O. Box 3038, Morogoro,
Tanzania

Denis Byamukama
Department of Biochemistry,
Makerere University, P.O. Box 7062, Kampala,
Uganda

Alexander K.T. Kirschner
Institute for Hygiene and Applied Immunology,
Medical University of Vienna,
Austria

Robert L. Mach
K. Brunner
Andreas H. Farnleitner (corresponding author)
Institute of Chemical Engineering,
Research Area Applied Biochemistry and Gene
    Technology,
Research Group Environmental Microbiology and
    Molecular Ecology,
Vienna University of Technology,
Gumpendorferstraße 1a, A-1060 Vienna,
Austria
E-mail: *A.FARNLEITNER@aon.at*

Alexander K.T. Kirschner
Andreas H. Farnleitner
InterUniversity Cooperation Centre Water and
    Health
www.waterandheatlh.at

INTRODUCTION

Inadequate maintenance of the microbial quality of groundwater in peri-urban areas of developing countries is compounded by many factors, including high-priced microbiological infrastructure and the low level of socioeconomic conditions (Butterfield & Camper 2004). As a result, peri-urban communities consume water of unknown quality, which may put them at an unacceptable risk of infection by enteric pathogens (Hutin *et al.* 2003). This situation calls for simple, feasible and easy-to-perform methods

to help peri-urban communities understand, react to and subsequently manage the quality of well water.

Sanitary inspection, which identifies actual and potential sources of contamination of groundwater abstraction points, was proposed by the World Health Organization (WHO 1997) as part of the comprehensive and complementary risk-based assessment of drinking water quality (WHO 2004; Luby *et al.* 2008). This proposal supports the operation and maintenance of water points by providing clear

guidance for remedial action to protect and improve the water supply (Luby *et al.* 2008). The WHO (1997) established a format for sanitary inspection forms consisting of a set of questions which have 'yes' or 'no' answers (cf. Supplementary Table S1, available online at http://www.iwaponline.com/jwh/010/117.pdf). The questions are structured such that 'yes' answers indicate that there is a reasonable risk of contamination and 'no' answers indicate that the particular risk appears to be negligible. Each 'yes' answer scores one point and each 'no' answer scores zero points. At the end of the inspection, the points are totalled, yielding a sanitary inspection risk score (in this study, referred to as a risk-of-contamination, or ROC, score). A higher ROC score represents a greater risk that drinking water is contaminated by faecal pollution from the area immediately surrounding the well (Lloyd & Batram 1991; WHO 1997; Godfrey *et al.* 2006; Luby *et al.* 2008; Vaccari *et al.* 2009; Parker *et al.* 2010).

Limited data exist on the performance of such sanitary inspection tools (Lloyd & Batram 1991; Godfrey *et al.* 2006; Luby *et al.* 2008; Vaccari *et al.* 2009; Parker *et al.* 2010), especially in most developing and tropical countries like Tanzania. Moreover, the few existing investigations are based on a small set of standard faecal indicator bacteria (SFIB) such as faecal coliforms, *Escherichia coli* and enterococci and are thus potentially prone to an indication bias or limited to give a more holistic picture of the actual faecal pollution situation (e.g. recent versus past faecal pollution). This is especially significant in tropical environments where SFIB may originate from non-faecal sources (such as soil) and proliferate in aquatic habitats under favourable conditions, thus resulting in detection at levels that may not reflect the actual extent of faecal contamination (Solo-Gabriele *et al.* 2000; Byamukama *et al.* 2005; Ishii & Sadowsky 2008). Consequently, there is a need for investigations concerning the performance of ROC scoring in tropical environments based on robust faecal pollution diagnostics.

The aim of this study was to evaluate the predictive capacity of ROC scoring, based on simple-to-perform sanitary inspections according to the WHO guidelines, regarding the extent of bacterial faecal pollution in well water in a peri-urban area of a tropical, developing country. The analysis was based on a selection of wells representing environments ranging from low to high presumptive faecal pollution risk. To integrate the information regarding the extent of bacterial faecal pollution, a multi-parametric data set including SFIB, alternative faecal indicators (i.e. *Clostridium perfringens* spores and sorbitol-fermenting Bifidobacteria) and physico-chemical parameters were investigated. We hypothesised that ROC scoring would accurately predict risk levels for bacterial faecal pollution in Tanzania, a tropical and developing area, thus representing a useful tool for well-water quality management.

MATERIALS AND METHODS

It is important to note that the study published by Mushi *et al.* in 2010 is based on the same field investigation described in this work. Mushi *et al.* (2010) focused on the ecology of sorbitol-fermenting Bifidobacteria (SFB) in streams and groundwater in a tropical environment, using lumped data from the collected well water samples. However, the current study focuses on the ROC scoring based on a detailed analysis of the same wells in a variety of locations and over a span of three months.

Selection of the wells

Wells were selected using weighted random sampling from the peri-urban area of Dar es Salaam, Tanzania (6.2° S, 39.2° E). The approach covered well environments ranging from low to high presumptive faecal pollution risk based on the stratified randomisation according to Byamukama *et al.* (2005). Wells W1, W2 and W3 are situated in heavily vegetated areas without nearby stagnant water, septic tanks, pit-latrines, human residences or anthropogenic activities. Residents travel approximately 200–300 m from their houses to collect water from these wells. Well W4 is situated approximately 100 m from human residences. Ponding was evident around the well. Neither septic tanks/pit-latrines nor fences were observed within 30 m of the well. Wells W5 and W6 were situated close (approximately 3–5 m) to roads. Septic tanks and pit-latrines were observed at approximately 15 m from the well. Ponding was evident at approximately 10 m from the well, as the

drainage channel was faulty. Although wells W1–W6 were well-constructed, no apparent protective measures had been taken against anthropogenic activities. Wells W7, W8 and W9 were chosen as they are situated in a highly populated area with poor sanitary infrastructure. These wells are poorly constructed and poorly protected. The walls of the wells are cracked, with broken platforms that allow growth of algae on the walls and influx of stormwater and runoff. The drainage channels are dirty and broken, allowing ponding. Some of the wells are close to roads and surface runoff, or are within 5 m of poorly constructed septic tanks/latrines situated uphill of the well site. A map of the well locations is given in Mushi *et al.* (2010).

Risk of contamination scoring

On-site sanitary inspections and well-water sampling from the nine chosen wells were performed from May to July 2005. This period was characterised by wet and dry climatic patterns. ROC scoring was performed according to the questions proposed by the WHO (1997) as given in Table S1 (supplementary material, http://www.iwaponline.com/jwh/010/117.pdf). The ROC scores range from a low risk of contamination (scores = 0–30%), through a medium (40–50%) or high (60–70%) risk of contamination, to a very high risk of contamination (80–100%). The method used to calculate these scores is described briefly above in the Introduction.

Water quality investigations

Water samples were taken at the time of the sanitary inspections from the nine wells described by Mushi *et al.* (2010) using 1 L sterile glass bottles and aseptic technique, according to American Public Health Association (APHA) standard methods (APHA 2000). Each well was sampled six different times over three consecutive months (from May to July 2005), yielding a total of 54 water samples. Field analysis (dissolved oxygen, temperature, electrical conductivity and salinity), microbiological analysis (presumptive total coliforms, TC; presumptive *E. coli*; presumptive faecal coliforms, FC; presumptive *C. perfringens*, CP; and presumptive SFB) and chemical parameters (nitrates plus nitrites, hardness and chlorides) were determined as described in Mushi *et al.* (2010).

Statistics

E. coli concentrations were expressed as logarithmic transformations ($\log_{10}$). Statistical analyses were performed with the Statistical Package for Social Sciences version 11.0 (SPSS Inc., Chicago, IL, USA). A non-parametric Kruskal–Wallis test was used to compare the median number of faecal bacteria in each contamination risk category. A Bonferroni correction was applied in cases where parameters were tested multiple times. The ROC score was calculated as the percentage of total positive answers to the ten questions posed by the checklist. A Spearman rank correlation analysis included the ROC score, potential microbial indicators of faecal pollution and physicochemical water quality parameters. A discriminant analysis of faecal indicator bacteria and ROC data, using Wilks' lambda and a within-group covariance matrix, was performed to cluster wells with the same risk level. Countable microbial colonies were expressed as colony forming units (CFU) per 100 mL. Probability (p) values below 0.05 were considered statistically significant.

RESULTS AND DISCUSSION

The surveyed wells showed a distinct pattern of ROC scores ranging from 20 to 100%. Based on the ROC scoring, the surveyed wells (W1–W9) could be placed into four categories, as suggested by the WHO (1997): low (W1–W3); medium (W4), high (W5 and W6) and very high (W7–W9) risk wells (Table 1). The ROC-based categories correlated remarkably well with the levels of bacterial faecal pollution as determined by the multi-parametric microbial pollution parameters (Table 1, $p < 0.05$; $n = 54$, Kruskal–Wallis Test). The ROC scores showed significant correlation with all indicators (*E. coli*, $r = 0.88$; TC, $r = 0.86$; FC, $r = 0.73$; CP = 0.67; $p < 0.05$, $n = 54$; see also the multiple correlation matrix [Table S2, supplemental material, http://www.iwaponline.com/jwh/010/117.pdf]) except for SFB ($r = 0.34$; $p > 0.05$, $n = 18$; Table S2). SFB could only be detected in water samples from very high risk wells (W7–W9), demonstrating that these wells were contaminated very recently with human faecal pollution (Mushi *et al.* 2010). It should be noted that the tight correlation between all of the investigated faecal indicators (except for the less abundantly

Table 1 | ROC scores (%), microbial indicators of faecal pollution (CFU/100 mL) and physicochemical parameters of water quality in peri-urban wells sampled during the study period

Wells	ROC (%)	Qualitative risk category[b]	Median (range) for[a]				
			TC	FC	*Escherichia coli*	*Clostridium perfringens*	SFB
W1	20	Low risk	19.0 (4–330)	11.0 (2–30)	2.0 (0–6)	1.7 (1–10)	nd
W2	30	Low risk	4.0 (2–10)	2.0 (1–10)	0.8 (0–6)	0.5 (0–8)	nd
W3	30	Low risk	19.0 (11–250)	8.0 (5–10)	5.0 (1–7)	0.3 (0–30)	nd
W4	50	Medium risk	35.0 (2–180)	16.0 (1–30)	7.5 (0–30)	26.0 (1.5–80)	nd
W5	60	High risk	75.0 (40–580)	32.0 (19–60)	11.0 (2–50)	37.5 (0–50)	nd
W6	70	High risk	65.0 (4–210)	24.0 (2–90)	19.5 (1–20)	0.0 (0–1)	nd
W7	80	Very high risk	49,500 (18,000–80,000)	6,200 (2,000–20,000)	2,800 (600–4,000)	800 (190–1,940)	400 (0–1,090)
W8	100	Very high risk	29,500 (2,900–68,000)	1,500 (600–6,500)	150 (0–1,300)	140 (32–310)	15.0 (0–1,200)
W9	100	Very high risk	92,500 (70,000–120,000)	6,500 (1,500–35,000)	6,000 (320–21,000)	1,200 (190–2,160)	400 (0–8,000)

Median (range) for

Temperature (°C)	Salinity (‰)	Chloride (mg/L)	Nitrate plus nitrite (mg/L)	Conductivity (μS)	Hardness (mg/L)	pH	DO (mg/L)
27.4 (26.9–28.9)	3.0 (2.9–3.0)	794 (550–852)	14 (7–66)	5,615 (5,500–5,920)	1,622 (778–1,825)	6.0 (5.8–6.8)	4.5 (1.1–7.8)
28.0 (26.9–29.8)	1.0 (0.6–1.1)	461 (370–486)	17 (1–49)	2,135 (1,068–2,229)	392 (381–450)	7.3 (7.2–7.5)	4.6 (1.9–8.4)
31.0 (29.0–33.6)	1.1 (0.5–1.3)	429 (296–515)	42 (12–92)	2,366 (1,066–2,869)	224 (213–262)	6.1 (6.1–6.2)	2.0 (1.3–2.5)
29.9 (28.5–32.7)	1.2 (0.6–1.3)	348 (285–362)	41 (19–51)	2,484 (1,330–2,669)	346 (189–364)	7.0 (6.7–7.1)	2.3 (1.5–3.7)
29.6 (27.8–29.9)	1.3 (0.6–1.3)	498 (371–537)	32 (9–59)	2,633 (1,186–2,807)	565 (336–661)	6.8 (6.7–6.9)	2.4 (1.5–7.3)
29.5 (28.4–31.1)	0.8 (0.4–0.8)	248 (220–259)	42 (19–75)	1,791 (794–1,857)	180 (176–213)	6.6 (6.5–6.8)	1.9 (1.3–2.8)
26.5 (25.8–29.3)	0.1 (0.1–0.1)	14 (9–28)	12 (3–24)	171 (156–245)	10 (10–45)	6.5 (6.3–6.8)	2.0 (0.5–8.1)
28.9 (28.0–29.3)	1.2 (0.8–1.3)	335 (265–391)	34 (8–104)	2,408 (1,259–2,689)	240 (192–310)	6.4 (6.3–6.4)	2.1 (1.5–6.3)
28.2 (27.7–29.2)	0.8 (0.7–0.8)	249 (142–287)	36 (11–76)	1,547 (1,494–1,786)	119 (60–199)	6.3 (6.1–6.9)	2.9 (1.9–6.1)

[a]Abbreviations: ROC, risk of contamination; TC, total coliforms; FC, faecal coliforms; SFB, sorbitol-fermenting Bifidobacteria; DO, dissolved oxygen; $n = 54$.
[b]Risk categorisation was performed according to World Health Organization (WHO 1997) criteria: 0–30% = low; 40–50% = medium; 60–70% = high; 80–100% = very high risk.

occurring SFB) is not unexpected, as all of the considered parameters are by definition associated with faecal contamination. The faecal indicator with the highest association with the ROC score was chosen to further investigate the predictive capacity of ROC scoring. Regression analysis indicated that the ROC score was able to successfully predict up to 87.4% of the *E. coli* concentrations among the investigated wells (Figure 1).

The results from a discriminant analysis of the multi-parametric bacterial pollution parameters and the ROC data, using Wilks' lambda and the within-group covariance matrix, closely mirrored the four clusters classified by ROC scoring (Figure 2). The first and second discriminants accounted for 96% of the observed variation, suggesting that the selected wells exhibit distinct levels of contamination and that the wells with the same level of contamination are tightly clustered by ROC scores and the bacterial faecal pollution data. In contrast, this tight grouping was not observed when ROC scores and physico-chemical parameters were subjected to discriminant analysis using the same algorithms because the first and the second discriminants accounted for <10% of the variation (see also Table S2 supplemental material online at http://www.iwaponline.com/jwh/010/117.pdf). This result is most likely due to comparable hydrogeochemical conditions among the investigated well catchments ($p > 0.05$, $n = 45–54$, Kruskal–Wallis test). These results provide
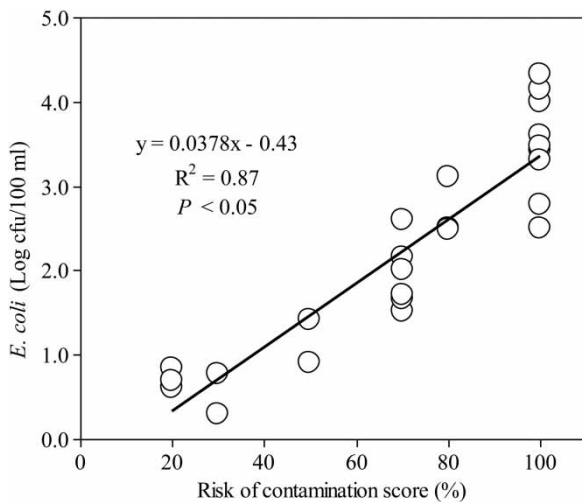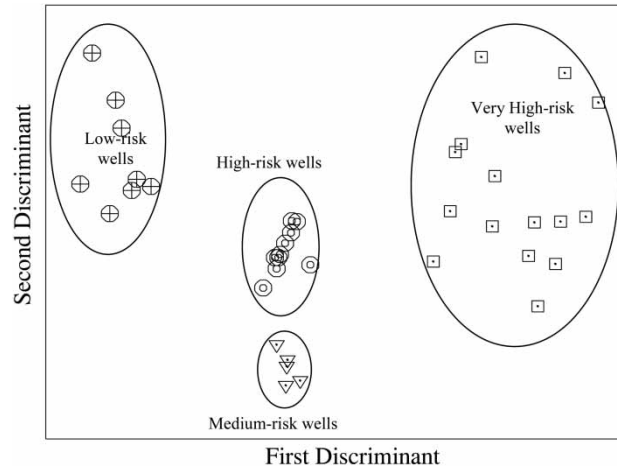

Figure 2 | A discriminant analysis of microbial indicators of faecal pollution and ROC scoring data, using Wilks' lambda and a within-group covariance matrix, cluster wells with the same level of contamination ($n = 18, 6, 12$ and 18 for wells with low, medium, high and very high ROC scores, respectively).

evidence that bacterial faecal pollution is closely linked to the sanitary hazards identified by the wells' ROC scores (Figure 1). Thus, the combination of sanitary inspection and analysis of microbiological water quality may also be useful for identifying the most important causes of and control measures for well contamination, which is important to support effective and rational decision-making. For instance, it will be important to know whether on-site (in association with the construction and maintenance of the location and its immediate inner protection zone) or off-site (in association with the catchment protection of the well) sanitation could be associated with the contamination of drinking water, as the remedial actions required to address either source of contamination will be very different. Such analysis may also identify other factors associated with contamination, such as heavy rainfall. It may be useful to complement such combined analysis with viral faecal indicators (e.g. somatic coliphages) to avoid underestimating viral mobility by focusing solely on bacterial parameters in porous groundwater resources. However, the results of the current study most likely relate equally to bacteriological, viral and protozoan faecal hazards, as microbial faecal contamination directly at the site of the well is considered to be the primary source of contamination (such as due to cracks in the well walls).

The high faecal pollution levels in various wells identified by the sanitary survey suggest that the well water is

Figure 1 | Regression analysis of the risk-of-contamination (ROC) score (%) and *Escherichia coli* ($\log_{10}$ CFU/100 mL) data.

not suitable for drinking without treatment (Mpanda 2002; Ntukula & Mosha 2002). As a result, appropriate remedial action such as fencing the well infrastructure, constructing standard well infrastructure, regular repair of the well infrastructure, positioning of latrines/septic tanks at a prescribed distance (>10 m) from wells and removal of ponding around the wells should be conducted to protect the well water. Furthermore, special informative and educational campaigns should be combined with the implementation of demonstrative pilot projects to improve the potential health impact in peri-urban communities (Vaccari *et al.* 2009).

The high ability of ROC scoring for predicting the extent of microbial faecal pollution, as reported in this study, appears to contradict results from previous investigations carried out in other tropical environments. In these studies, ROC scores did not correlate with any of the assessed indicators (Luby *et al.* 2008; Parker *et al.* 2010). However, it is worth noting that only single-parameter based standard faecal indicators were used, and thus, the approach might not have reflected the actual extent of faecal pollution (Luby *et al.* 2008; Parker *et al.* 2010). This indication bias did not occur in our study, as shown by the consensus of results based on multiple indicators, and other factors probably led to the results observed. The current study is also distinguished by a selection of wells spanning a range of environments from low to high presumptive faecal pollution risk. Indeed, the wide range of faecal pollution levels detected in the well water samples, from very low (i.e. 0–5 *E. coli* CFU per 100 mL) to extensive (i.e. 21,000 *E. coli* CFU per 100 mL), likely supported statistical analysis to identify associated factors. Analysis of the ROC scoring matrix results led to another important finding in the current study, specifically, a clear association ($r = 0.67$; $p < 0.05$) between the extent of faecal pollution risk (sum of scores from questions 1 and 2) with failures in technical well infrastructure and protection (sum of scores from questions 3–10) was discernable. Although well selection was based on the criteria of presumptive faecal pollution risk, increased faecal pollution clearly correlated with decreased technical well infrastructure and protection in the environments investigated in this study. This finding is not surprising and could possibly be generalised to other, similar regions. The ROC scoring evaluation in this study was thus tested in 'bad' to 'worst case' scenarios in terms of faecal pollution risk and the technical infrastructure and protection. The ROC scoring matrix reflected this situation remarkably well, indicating its robust nature.

It should be noted that determining ROC scores is a simple procedure as it requires minimal training in terms of the identification of potential sanitary hazards, calculating the ROC score and interpreting the score with the aid of the conditions highlighted by the WHO (1997). Moreover, compared with more advanced methods, ROC scoring is less expensive, produces results about the actual and potential sources of contamination within a very short period of time, and can be used to select appropriate remedial actions to improve well water quality. One potential limitation is that the ROC score only reflects the environment immediately surrounding the well. Groundwater contamination further away from the well installation is not considered, and may therefore lead to water quality problems if the aquifer lacks the capacity for self-purification. For example, fractured aquifers in karst areas may cause rapid contamination of the well water. Sanitary inspection methods such as ROC scoring thus represent only one element in the chain of water quality management and cannot replace basic hydrogeological investigations.

CONCLUSIONS

Sanitary inspection of wells using ROC scoring – according to established WHO procedures – indicated a remarkable ability to predict the extent of bacterial faecal pollution of the wells tested in this study. ROC scoring appears to be a useful and cost-effective tool for supporting well-water management at these locations of the tropical peri-urban environment of Tanzania. However, the selected region and its associated conditions may have increased the apparent relationship between the ROC scores and the faecal pollution levels observed in the well-water samples. The predictive performance of ROC scores should thus only be extrapolated to other tropical regions with caution. Nevertheless, the present data can be taken as a promising start. Further research in other geographical regions should focus on a detailed determination of the chemical, physical and microbiological aspects of the water quality; a selection

of wells representing environments with differing faecal pollution risks and varying degrees of technical well infrastructure; and protection efforts in the selected environments. Furthermore, the ROC scoring procedure could be further developed to yield more detailed information about the pollution risk as well as the technical well infrastructure and protection efforts. In this regard, analysis of the multi-parametric data as well as the selection of the most informative parameters could be performed using alternative statistical and mathematical tools. Machine learning tools such as artificial neural networks or genetic programming could help improve the analysis (e.g. by detecting non-linear relationships in the data matrix) and enhance the prediction of faecal pollution and risk (Chau *et al.* 2002; Muttil & Chau 2006; Muttil & Chau 2007; Winter *et al.* 2007).

Microbiological pollution levels exceeded the WHO drinking water quality recommendations throughout the investigation period. In particular, high risk and very high risk category wells need improved construction to block contamination pathways, improved hygiene practices around the well and comprehensive management measures to protect wells from anthropogenic activities. Furthermore, water from high to very high risk categories should undergo further treatment such as chlorination or boiling to meet the microbiological standard for drinking water. Very high risk category wells in urban areas were demonstrably contaminated by very recent human faecal materials, as revealed by the human-associated source-tracking parameter of SFB. These wells could pose a critical factor in the transmission of waterborne pathogens between the human populations.

ROC scoring may also be a useful tool for selecting representative sites or identifying hotspots for epidemiological surveys or quantitative microbial risk assessment (QMRA) studies. However, it should be emphasised that a correlation between the risk of infection and the extent of faecal pollution can only be expected when there is concurrent shedding of faecal indicators and pathogens. Detailed quantitative information regarding the occurrence of carefully selected reference pathogens is required. We have thus refrained from any attempt to translate faecal indicator levels into levels of infection risk with non-site-specific, and therefore unreasonable, pathogen-to-indicator ratios in this study.

ACKNOWLEDGEMENTS

REFERENCES

APHA 2000 *Standard Methods for the Examination of Water and Wastewater*, 20th edition. American Public Health Association/American Water Works Association/Water Environmental Federation, Washington, DC.

Butterfield, P. W. & Camper, A. K. 2004 Development of a toolbox to assess microbial contamination risks in small water systems. *J. Wat. Health* **2** (4), 217–232.

Byamukama, D., Mach, R. L., Kansiime, F., Manafi, M. & Farnleitner, A. H. 2005 Discrimination efficacy of faecal pollution detection in different aquatic habitats of a high altitude tropical country using presumptive coliform, *Escherichia coli* and *Clostridium perfringens* spores. *Appl. Environ. Microbiol.* **71** (1), 65–71.

Chau, K. W., Cheng, C. T. & Li, C. W. 2002 Knowledge management system on flow and water quality modeling. *Expert Syst. Appl.* **22** (4), 321–330.

Godfrey, S., Timo, F. & Smith, M. 2006 Microbiological risk assessment and management of shallow groundwater sources in Lichinga. *Mozambique Wat. Environ. J.* **20**, 194–202.

Hutin, Y., Luby, S. & Paquet, C. 2003 A large cholera outbreak in Kano city, Nigeria: the importance of hand washing with soap and the danger of street-vended water. *J. Wat. Health.* **1** (1), 45–52.

Ishii, S. & Sadowsky, M. J. 2008 *Escherichia coli* in the environment: implications for water quality. *Microbes Environ.* **23** (2), 101–108.

Lloyd, B. J. & Batram, J. 1991 Surveillance solutions to microbiology problems in water quality control in developing countries. *Wat. Sci. Tech.* **24**, 61–75.

Luby, S. P., Gupta, S. K., Sheikh, M. A., Johnston, R. B., Ram, P. K. & Islam, M. S. 2008 Tubewell water quality and predictors of contamination in three flood-prone areas in Bangladesh. *Appl. Microbiol.* **105**, 1002–1008.

Mpanda, S. 2002 Groundwater pollution in Dar es Salaam city. In: *Current Environmental Aspects in Dar es Salaam, Tanzania*

(A. K. Kivaisi & Y. D. Mgaya, eds), University Press, Dar es Salaam, pp. 85–93.

Muttil, N. & Chau, K. W. 2007 Machine learning paradigms for selecting ecologically significant input variables. *Eng. Appl. Artificial Intel.* **20** (6), 735–744.

Muttil, N. & Chau, K. W. 2006 Neural network and genetic programming for modelling coastal algal blooms. *Int. J. Environ. Pollut.* **28** (3–4), 223–238.

Mushi, D., Byamukama, D., Kivaisi, A. K., Mach, R. L. & Farnleitner, A. H. 2010 Sorbitol-fermenting Bifidobacteria are indicators of very recent human faecal pollution in streams and groundwater habitats in urban tropical lowlands. *J. Wat. Health* **8** (3), 466–478.

Ntukula, J. W. & Mosha, D. M. S. 2002 Water supply and quality problems in squatter areas of Dar es Salaam city, Tanzania. In: *Current Environmental Aspects in Dar es Salaam, Tanzania* (A. K. Kivaisi & Y. D. Mgaya, eds), University Press, Dar es Salaam, pp. 104–111.

Parker, A. H., Youlten, R., Dillon, M., Nussbaumer, T., Carter, R. C., Tyrrel, S. F. & Webster, J. 2010 An assessment of microbiological water quality of six water source categories in north-east Uganda. *J. Wat. Health* **8** (3), 550–560.

Solo-Gabriele, H. M., Wolfert, M. A., Desmarais, T. R. & Palmer, C. J. 2000 Sources of *Escherichia coli* in a coastal subtropical environment. *Appl. Environ. Microbiol.* **66**, 230–237.

Vaccari, M., Collivignarelli, C., Tharnpoophasiam, P. & Vitali, F. 2009 Well sanitary inspection and water quality monitoring in Ban Nam Khem (Thailand) 30 months after 2004 Indian Ocean Tsunami. *Environ. Monit. Assess.* **161** (1–4), 123–133.

Winter, C., Hein, T., Kavka, G., Mach, R. L. & Farnleitner, A. H. 2007 Longitudinal changes in the bacterial community composition of the Danube River: A whole river approach. *Appl. Environ. Microbiol.* **73**, 421–431.

WHO 1997 *Guidelines for Drinking Water Quality, Surveillance and Control of Community Supplies*, 2nd edition, vol. 3. World Health Organization, Geneva, Switzerland.

WHO 2004 *Guidelines for Drinking Water Quality: Recommendations*, 3rd edition, vol. 3. World Health Organization, Geneva, Switzerland.