35th WEDC International Conference, Loughborough, UK, 2011

THE FUTURE OF WATER, SANITATION AND HYGIENE:
INNOVATION, ADAPTATION AND ENGAGEMENT IN A CHANGING WORLD

Evaluation of different strategies for deploying the H2S test to detect microbial contamination of drinking water

J. A. Wright, H. Yang, S. Pedley & S. Gundry, UK

REFEREED PAPER 1135

The hydrogen sulphide (H_2S) test has been proposed as a presence-absence, low-cost field test to detect microbial contamination for nearly 30 years and is now widely used in many countries. The objective of this study is to identify different strategies for deploying the test and assess how each might be affected by the test's accuracy. Evidence on the H_2S test's accuracy is drawn from a recent systematic review. We identified six different strategies for deploying the test based on a literature review. Three strategies used the H_2S test in isolation, while the other three used the H_2S test in combination with standard, laboratory-based methods or alongside sanitary risk inspection surveys. We conclude that using the test in combination with laboratory-based methods or sanitary risk inspections reduces the problems posed by false positive H_2S test results. However, such strategies may be more costly and complex to implement.

Introduction

Water quality monitoring may be difficult to implement in remote or resource-poor settings, which lack the necessary infra-structure, finances, and trained personnel. One solution that has been proposed to this problem is the use of low-cost, field tests for drinking water quality (Cherukuri and Anjaneyulu, 2005; Manja et al., 1982). In this study, we examine the strengths and weaknesses of different strategies for deploying one such test, the Hydrogen Sulphide (H_2S) test for microbial contamination of drinking water. To illustrate the impact of field test diagnostic accuracy on each of these strategies, we draw on a recent systematic review of the diagnostic accuracy of the H_2S test.

Manja et al. (1982) first proposed the H_2S method. Their original formulation contained 20g of peptone, 1.5g of dipotassium hydrogen-phosphate, 0.75g of ferric ammonium citrate, 1g of sodium thiosulphate, 1ml of teepol and 50ml of water. 1 ml of the medium was then absorbed onto a strip, placed in a sample bottle, sterilised, and dried for subsequent use. After collection, water samples are poured into the bottle, shaken, and then incubated for a period of at least half a day, often under ambient conditions, before being examined for a possible colour change. After incubation, a change to a black or grey colour indicates the presence of H_2S producing bacteria, which are then taken as an indicator of faecal contamination of the water. Since the method involves a single 20ml sample bottle, it is implemented as a presence/absence test. The H_2S method is cheaper than recognized methods and does not require laboratory facilities.

H2S diagnostic accuracy

Much research has focused on the performance of low-cost field tests such as the H_2S method (e.g. Gupta et al., 2008; Sobsey and Pfaender, 2003). Typically, this research has compared H_2S test presence/absence results with those of standard, laboratory-based methods based on the recognized indicator organisms, thermotolerant coliforms or *E. coli*. By comparing results, four groups of samples can be identified (see Table 1):

- True positives (tp), where both tests suggest that water is faecally contaminated;
- False positives (fp), where the field test is positive but a standard, laboratory-based method is negative.
- True negatives (tn), where both tests suggest no evidence of faecal contamination.

- False negatives (fn), where the field test is negative, but a laboratory-based method suggests the sample is faecally contaminated.

Table 1: The calculation of measures of field test diagnostic accuracy

	***E. coli* or thermotolerant coliform test result**	
H_2S test result	**Positive**	**Negative**
Present	true positive (tp)	false positive (fp)
Absent	false negative (fn)	true negative (tn)

From these four groups of samples, sensitivity and specificity are commonly used by researchers to measure test accuracy. Sensitivity (=tp/(tp+fn)) is the proportion of true positives (faecally polluted water samples) that are correctly identified by the field test (Altman and Bland, 1994). Specificity (=tn/(tn+fp)) is the proportion of true negatives (safe water according to the recognized microbiological standard) that are correctly identified by the field test (Altman and Bland, 1994).

Recently, we reviewed literature comparing the H_2S test with standard methods (Wright et al., under review), drawing on 13853 water samples collected in 51 different studies. We estimated the overall sensitivity of the H_2S test to be 0.88 (95% confidence limits 0.77 to 0.94) relative to *E. coli* and 0.87 (0.80 to 0.92) relative to thermotolerant coliforms. We estimated its specificity to be 0.75 (0.55 to 0.88) relative to *E. coli* and 0.82 (0.72~0.90) relative to thermotolerant coliforms. We also found significant variation in the test's accuracy between studies, which proved difficult to explain using testing procedures, water sample or study characteristics. For example, a South African study of the H_2S test (Genthe and Franck, 1999) estimated its sensitivity as 0.93 (95% confidence limits 0.89 to 0.96) relative to thermotolerant coliforms. A similar study in Thailand (Hewison et al., 1988) found the test's sensitivity to be just 0.60 (confidence limits 0.46 to 0.72) relative to thermotolerant coliforms and a study in Burkina Faso estimated the test's sensitivity as 0.27 (0.06 – 0.61) (Monjour et al., 1986). To understand the practical implications of such findings, we now look at the way such test results are used in practice.

Strategies for deploying the H_2S test in the field

Based on a further literature review, we identified six different strategies for deploying the H_2S test: three where the field test is used independently of laboratory testing, and three where the field test is used alongside a conventional, laboratory-based test or with sanitary risk inspection.

H_2S testing without standard laboratory testing

1. The most straightforward monitoring situation where **a field test can be deployed** where a single sample is taken, processed using the field test, and an operational decision reached about a particular water source.
2. The slightly more complex operational monitoring situation where **a small number (typically two but occasionally more) of field tests** are performed and their results are then compared. As one example, a field water test might be used to test a supply before an intervention takes place. Following the testing of this baseline sample, a follow-up field test is used to check that the initial intervention has worked (Pumphrey et al., 2006).
3. The situation where **groups of field test results** are compared, for example to understand the relative levels of contamination across different areas, over different periods, or across different supply types. In a survey of the water sources on a remote Haitian island, for example, one study reported failure rates for different source types: '*Tests for the presence of hydrogen sulfide-producing bacteria were negative for samples collected from six out of seven drilled wells that were tested…. six out of eight capped springs that were sampled tested positive for the presence of hydrogen sulfide-producing bacteria*' (Troester and Turvey, 2004).

H_2S testing with laboratory testing or sanitary risk inspection

1. The field test is used as a **screening tool**, with a standard, laboratory-based testing being conducted as a subsequent follow-up if a sample is positive (UNICEF, 2007). For example, Mosley and Sharp (2005)

suggest: *'If a positive result is observed, another sample can be collected for further analysis by conventional means e.g. for faecal coliform enumeration.'*

2. The field test is put into widespread use in remote locations, but a **subset of water samples are processed in duplicate using a standard method** to evaluate field test performance. There are documented examples of such usage for the H_2S test, such as a survey undertaken in Nepal (Joshi and Maharjan, 2003): '*Over 150 different water sources were monitored, including traditional community taps (stone taps), household connections, shallow wells, deep wells, and household water storage tanks, using a low-cost bacterial test (H_2S) prepared locally. Ten percent of sample replicates were also tested for Total Coliforms and E. coli as quality control*.'
3. **H_2S testing is combined with sanitary risk inspection surveys** and decisions about source management consider both risk inspection scores and test results together. Such an approach has been advocated in Fiji (Mosley and Sharp, 2005). *'H_2S test results must be considered in parallel with the results of a sanitary survey…… For example, if a drinking water well is unprotected and the results of the H_2S test are positive on the first day, the users should be informed that a risk to health is likely, and steps must be taken to disinfect the water.'*

Assessment of strategies for deploying the H_2S test

Scenario 1: Simple use of a single H_2S test

Our earlier review (Wright et al., under review) suggests that the H_2S test generally provides an approximate indicator of faecal contamination. For example, its overall sensitivity was 0.88 compared to *E. coli* test results. In other words, the test detected 88% of samples contaminated with *E. coli*, so most but not all contaminated sources will be detected. Some will be missed because of false negatives. Furthermore, the overall summary figure of 88% sensitivity with *E. coli* hides much variation between studies. In one quite large study (Desmarchelier *et al*, 1992), for example, H_2S test sensitivity was just 0.62 relative to *E. coli*, well below the overall sensitivity estimate of 0.88.

Aside from false negatives, a false positive field test result could also have undesirable consequences, such as boiling drinking water unnecessarily, switching to more distant water sources, or losing faith in a field test's reliability. The proportion of times a positive H_2S test result proves to be a 'false alarm' (its positive predictive value) does not depend solely on the accuracy of the test. It also depends on how widespread water source contamination is. Where contaminated water sources are commonplace, positive H_2S results will frequently reflect genuine contamination. Where contaminated water sources are rare, the small number of true positive H_2S results will be outnumbered by false positives, producing many more 'false alarms'. In this situation, the harm done by false positives may outweigh the benefits of testing.

In summary, as shown in Table 2 for this scenario, the H_2S test does seem to provide useful information about source contamination, but this information needs to be balanced against the harm of false positives, particularly when source contamination rates are quite low. Furthermore, there are a minority of situations where the H_2S test performs poorly, suggesting that it needs to be benchmarked against laboratory methods before being deployed in a setting for the first time.

Scenario 2: 'Before' and 'after' H_2S tests

The issues noted for Scenario 1 above also apply to this scenario. In addition, a presence / absence test like H_2S will not always detect incremental improvements in water quality, where bacteria counts are reduced but remain detectable. The appropriateness of low-cost testing in this scenario partly also depends on how far false positives or false negatives recur again and again for the same source. If the same source repeatedly produces false positives, for example, then it will be unclear whether a treatment intervention has worked. There is currently no evidence to show whether the same sources repeatedly produce false positive H_2S results or whether false positives occur randomly, affecting first one source and then another. Without knowing more about how far misleading H_2S results recur, it is difficult to evaluate this use of the test.

Scenario 3: Comparing H_2S positive rates across areas / sources

Given that the H_2S test's diagnostic accuracy varies between settings (Wright et al., under review), it follows that it is impossible to understand whether a high contamination rate for an area really reflects a high false positive rate or a genuinely high rate of contamination. In addition, the use of field test results in this scenario is particularly problematic if source type affects the sensitivity and specificity of the field test. Often, the mix of water sources will vary by area and if sensitivity and specificity are affected by source type, then this will bias estimated contamination rates for different areas. For example, it has been suggested that the H_2S test may give more false positive results for groundwater than surface water samples, because

of the presence of H_2S-producing bacteria of non-faecal origin (Sobsey and Pfaender, 2003). If such differences in field test diagnostic accuracy do exist, then the contamination rate for an area with a comparatively high proportion of groundwater sources will be inflated, relative to other areas surveyed. It therefore seems difficult to justify its use in this way.

Scenario 4: H_2S as a screening test

Screening – where a positive H_2S test is followed up with a standard laboratory test – reduces the problem of false positives described above in Scenario 1. This strategy is particularly well suited to tests with high sensitivity, but lower specificity. Our review suggested that overall relative to *E. coli*, the H_2S test had higher sensitivity (0.88) than specificity (0.75), making it a good candidate screening test. However, low specificity still has implications in this scenario, since as specificity decreases, there is an increased frequency in unnecessary, more expensive laboratory-based testing. There is also a potential additional complication under this scenario. Where a supply system experiences brief contamination events or 'spikes' (e.g. following rainfall events or water pressure changes within a system), the contamination event may have passed by the time any follow-up sample is taken and the laboratory-based test may generate a negative result.

Scenario 5: H_2S tests supported by limited laboratory testing

As noted earlier in Scenario 1, there is wide variation in the diagnostic accuracy of the H_2S test in different studies. The approach outlined in this scenario could be a useful first phase in deploying the test, enabling its diagnostic performance in a particular context to be evaluated prior to use (e.g. as a screening tool as per Scenario 4, or in isolation as per Scenarios 1 to 3). There may also be value in having an ongoing programme through which a proportion of low-cost tests are also tested using laboratory-based methods. However, as with the scenario 4, standard laboratory-based testing will increase costs, monitoring complexity and be impossible in the remotest areas.

Scenario 6: H_2S testing with sanitary risk inspection

In this strategy, decisions around source management are based on H_2S test results and sanitary risk inspection scores. A source with a poor sanitary risk inspection score and a positive H_2S test result will therefore be considered a higher priority for remediation than one which has a positive H_2S test but a good sanitary risk inspection score. Since test results are considered alongside risk inspection scores, this strategy should therefore somewhat offset the consequences of any false positive and false negative results. In contrast, sanitary risk inspection will require more time and training, adding to monitoring complexity.

Table 2 summarises our evaluation of the six different strategies. Another option for deploying the H_2S test would be to undertake multiple tests of the same water sample, in effect changing it from a presence-absence test for contamination to a multiple tube fermentation test. Using more than one H_2S test on water drawn from the same sample enables the level of contamination to be quantified as a most probable number (in cfu/100ml), rather than simply recording presence or absence of contamination. Although several research studies have undertaken H_2S testing in this way (e.g. Kromoredjo and Fujioka, 1991), we were unable to identify any reports of multiple tube H_2S testing occurring in routine monitoring. This may be because performing and interpreting multiple H_2S tests adds to the complexity of the technique and increases the training requirements needed before deploying it.

Table 2. Summary of the strengths and weaknesses of each strategy for H_2S test field use

Strategies	Strengths	Weaknesses
1: Simple use of a single H_2S test	Simple to implement	False positives may outnumber true negatives if most sources are contamination-free.
2: 'Before' and 'after' H_2S tests	Simple to implement	Unclear how often false positives are repeatedly associated with same source; presence/absence testing may not detect incremental water quality improvements

3: Comparing H_2S contamination rates across areas / sources		H_2S test accuracy may vary from place to place
4: H_2S as a screening test	Reduces effects of false positives	More costly and more complex to implement; not possible in remote areas; brief contamination 'spikes' may be missed
5: H_2S testing supported by limited laboratory testing	Quantifies accuracy of H_2S test	More costly and more complex to implement; not possible in remote areas
6: H_2S testing with sanitary risk inspection	Reduces effects of false positives and negatives	More complex to implement

Conclusions

In general, combining the H_2S test with other information sources – whether laboratory testing or sanitary risk inspection surveys (Strategies 4 to 6) – offsets some of the consequences of limited test accuracy. Our systematic review suggests the H_2S test has quite high sensitivity, at least in some settings. This suggests it has value where the alternative is no microbiological testing of water quality and where contaminated water sources are found quite frequently. Particularly in less remote areas where it has been shown to have especially high sensitivity, it could prove a useful screening tool when used with standard laboratory-based methods (as per Scenario 4). However, its diagnostic accuracy does vary considerably between settings, making it inadvisable to compare contamination rates across areas (Scenario 3). In a situation where the H_2S test has never previously been evaluated, it therefore seems prudent to evaluate its diagnostic accuracy before use. This can be done by testing a random sample of water samples in duplicate using both H_2S and a standard, recognised test for *E. coli* or thermotolerant coliforms (as per Scenario 5).

The review methodology applied to the H_2S here could also potentially be adapted to look at the accuracy of other simple methods for assessing the risk of faecal contamination in the field, such as sanitary risk inspection checklists (WHO, 2002).

Acknowledgements

This research has been funded by the Bill & Melinda Gates Foundation, under grant reference 48599 ('Aquatest 2: Delivery a low cost water test for developing countries').

References

Altman, D.G. and Bland, J.M. (1994) *Diagnostic tests. 1: Sensitivity and specificity*. British Medical Journal Vol. 308, No 6943, pp. 1552.

Cherukuri, J. and Anjaneyulu, Y. (2005) *Design and development of low cost, simple, rapid and safe, modified field kits for the visual detection and determination of arsenic in drinking water samples*. International Journal of Environmental Research and Public Health Vol. 2, No 2, pp. 322-327.

Desmarchelier, P., Lew, A., Caique, W., Knight, S., Toodayan, W., Isa, A.R., and Barnes, A. (1992) *An evaluation of the hydrogen sulfide water screening-test and coliform counts for water-quality assessment in rural Malaysia*. Transactions of the Royal Society of Tropical Medicine and Hygiene Vol. 86, No 4, pp. 448-450.

Genthe, B. and Franck, M. (1999) *A tool for addressing microbial water quality in small community water supplies: an H_2S strip test*. Water Research Commission Report No. 961/1/99: Pretoria, South Africa.

Gupta, S.K., Sheikh, M.A., Islam, M.S., Rahman, K.S., Jahan, N., Rahman, M.M., Hoekstra, R.M., Johnston, R., Ram, P.K. and Luby, S. (2008) *Usefulness of the hydrogen sulfide test for assessment of water quality in Bangladesh*. Journal of Applied Microbiology Vol. 104, No 2, pp. 388-395.

Hewison, K., Mack, K. F. and Sivaborvorn, K. (1988) *Evaluation of a hydrogen sulphide screening test*. AIDAB, Thai-Australian Northeast Village Water Resource Project, report no 47: Melbourne, Australia.

Joshi, D. and Maharjan, M. (2003) "Urban ecosystems and health in Kathmandu: Community-based biological assessment of drinking water sources", in Bunch, M. J, Suresh, V. M, and Kumaran, T. V.

(ed.) *Proceedings of the Third International Conference on Environment and Health*.University of Madras and York University: Chennai, India, pp. 186-195.

Kromoredjo, P. and Fujioka, R.S. (1991) *Evaluating 3 simple methods to assess the microbial quality of drinking-water in Indonesia*. Environmental Toxicology and Water Quality Vol. 6, No 2, pp. 259-270.

Manja, K.S., Maurya, M.S. and Rao, K.M. (1982) *A simple field test for the detection of faecal pollution in drinking water*. Bulletin of the World Health Organization Vol. 60, No 5, pp. 797-801.

Monjour, L., Henry, P. and Guillemin, F. (1986) *Comparative-study of field-tests for the detection of polluted drinking-water in Sahel Regions*. Transactions of the Royal Society of Tropical Medicine and Hygiene Vol. 80, No 1, pp. 172-173.

Mosley, L.M. and Sharp, D.S. (2005) *The hydrogen sulphide (H_2S) paper-strip test: a simple test for monitoring drinking water quality in the Pacific Islands*. SOPAC Technical Report 373: Suva, Fiji Islands.

Pumphrey, S.I., Divelbiss, D.W. and Oerther, D.B. (2006) "Appropriate technology for the treatment of drinking water in Roche, Tanzania", in Gimbel,R., Graham,N., and Collins,M.R. (ed.) *Recent Progress in Slow Sand and Alternative Biofiltration Processes.* IWA Publishing: London, pp. 283-290.

Sobsey, M. D. and Pfaender, F. K. (2003) *Evaluation of the H_2S method for detection of fecal contamination of drinking water*. World Health Organisation: Geneve.

Troester, J.W. and Turvey, M.D. (2004) *Water-resources reconnaissance of Ile de la Gonave, Haiti*. Hydrogeology Journal Vol. 12, No 2, pp. 224-236.

UNICEF (2007) *UNICEF Handbook on water quality*. UNICEF: New York.

WHO (2002) *Guidelines for drinking-water quality (2nd Edition). Volume 3: Surveillance and control of community supplies*. World Health Organization: Geneva.

Wright, J.A., Yang, H., Walker, K., Pedley, S., Elliott, J. and Gundry, S.W. (under review) *A systematic review of the diagnostic accuracy of the hydrogen sulphide test versus indicator bacteria tests for faecal contamination of water*. Water Research.

Keywords

H_2S method, sensitivity, specificity, faecal pollution, water safety plans

Contact details

Jim A Wright
School of Geography, University of Southampton, Highfield, Southampton SO17 1BJ, UK.
Tel:+44 (0)2380-594619
Fax:+44 (0)2380-593295
Email: j.a.wright@soton.ac.uk
http://www.southampton.ac.uk/geography/staff_profiles/academic/jaw3.html

Hong Yang
School of Geography, University of Southampton, Highfield, Southampton SO17 1BJ, UK.
Tel: +44 (0)2380-592253
Fax: +44 (0)2380-593295
Email: h.yang@soton.ac.uk
http://www.southampton.ac.uk/geography/staff_profiles/research/hy5g09.html

Steve Pedley
Robens Centre for Public and Environmental Health, University of Surrey, Guildford, Surrey GU2 7XH, UK
Tel: +44 (0)1483 689281
Email: S.Pedley@surrey.ac.uk
http://www.rcpeh.com/index.php?option=com_content&task=view&id=102&Itemid=242

Stephen Gundry
Water and Health Research Centre, Dept of Civil Engineering, University of Bristol, Merchant Venturers Building, Woodland Road, Clifton, Bristol BS8 1UB, UK.
Tel: +44 (0)117 331 5296
Fax: +44 (0)117 331 5296
Email: stephen.gundry@bristol.ac.uk
http://www.bristol.ac.uk/civilengineering/person/details.html?id=36710