

Sensitivity and Specificity Analysis Relation to Statistical Hypothesis Testing and Its Errors: Application to *Cryptosporidium* Detection Techniques

Emmanuel de-Graft Johnson Owusu-Ansah¹, Angelina Sampson², Amponsah K. Samuel³, Abaidoo Robert⁴

¹Department of Mathematics and Statistics, University of Energy and Natural Resources, Sunyani, Ghana

²Department of Chemistry and Biochemistry, South Dakota State University, Brookings, USA

³Department of Mathematics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

⁴Department of Theoretical and Applied Biology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Email: degraftt@gmail.com, emmanuel.owusu-ansah@uenr.edu.gh, vaishalu788@gmail.com, skamponsah.cos@knust.edu.gh, abaidoorc@yahoo.com

Received 4 February 2016; accepted 11 April 2016; published 14 April 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The use of Statistical Hypothesis Testing procedure to determine type I and type II errors was linked to the measurement of sensitivity and specificity in clinical trial test and experimental pathogen detection techniques. A theoretical analysis of establishing these types of errors was made and compared to determination of False Positive, False Negative, True Positive and True Negative. Experimental laboratory detection methods used to detect *Cryptosporidium* spp. were used to highlight the relationship between hypothesis testing, sensitivity, specificity and predicted values. The study finds that, sensitivity and specificity for the two laboratory methods used for *Cryptosporidium* detection were low hence lowering the probability of detecting a “false null hypothesis” for the presence of cryptosporidium in the water samples using either Microscopic or PCR. Nevertheless, both procedures for cryptosporidium detection had higher “true negatives” increasing its probability of failing to reject a “true null hypothesis” with specificity of 1.00 for both Microscopic and PCR laboratory detection methods.

Keywords

Sensitivity, Specificity, Type I and II Errors, Detection Methods, Hypothesis Testing

1. Introduction

Measuring the effectiveness of procedures and methods has over the years received vast consideration. It's worth noting that, a diagnostic test serves as a guide to physicians in assessment of diseases, and a statistical inference theory also serves as a guide to scientist in testing hypothesis. Mostly, there is a thin line of difference between clinical diagnostic testing and hypothesis testing; however clinicians are more familiar with diagnostic than hypothesis and vice versa [1].

Sensitivity and specificity have been used for diagnostic accuracy most often in enumeration process [2]. Moreover, these terms are mostly used in medical research as well as biological research. Sensitivity and specificity play a major role in estimating the efficiency of scientific procedural stages used in carrying out research through either empirical approach or experimental approach. In biological settings, it helps to estimate the recovery rates for pathogen detections. These terms have been seen as an independent for scientific procedural measurements for experiments; however its links to the measurement of the Null Hypothesis Testing Procedure (NHTP) give the backbone of statistical interpretation of hypothesis testing results and the needed theoretical background. In evaluating any new diagnostic test before introducing into clinical settings or evaluation of new protocol in laboratory experiment, most of these new methods are evaluated in relation to a benchmark which is mostly an old previously accepted and historically reliable gold standard and hence need the theoretical backbone of statistical interpretation.

2. Sensitivity and Specificity

In clinical and laboratory test, the imperative answers seeking is whether the accepted protocol or diagnostic procedure test is either sensitive enough to detect the presence of a disease/pathogen in a contaminated sample or is the test specific enough to indicate the absence of a diseases or pathogen in samples which are in fact not contaminated with the pathogen. Primarily, definition of sensitivity is the probability measuring the likelihood for a test to pick up the presence of a disease/pathogen, alternatively, a true positive is recorded when a procedure reflects the presence of pathogen in a contaminated sample. Furthermore, we define specificity as the probability of measuring the likelihood for a test to pick up the absence of a disease/pathogen, alternatively, a true negative is recorded when a procedure reflects the absence of a pathogen when the sample is not contaminated [1]. In contrast, a false positive occurs when the test reports a positive result for a person who is disease free or a positive result for a pathogen free sample, whereas a false negative occurs when the test reports a negative result for a person who actually has the disease or negative result for a pathogen infested sample [3].

2.1. Hypothesis Testing

Hypothesis is a numerical statement of an unknown parameter. Reference [4] defined hypothesis testing as a scientific assertion that is testable on the basis of observing a process that is modeled via a set of random variables. In hypothesis testing, there is a “null hypothesis” which corresponds to a presumed default “state of nature”. Corresponding to the null hypothesis is an “alternative hypothesis” which relates to the opposition situation.

2.2. Type I, Type II Errors and Predictive Value

When a test is conducted and the resulting test does not match the true state of the condition, then an error has occurred. The two kinds of error in statistical hypothesis error testing are “Type I and Type II errors” depending upon which hypothesis has incorrectly been identified as the true state of nature.

Type I occurs when the statistical test falsely indicates otherwise when a true state exist, it is analogues “false positive” in diagnostic test, on the other hand, Type II occurs when statistical test fails to recognize the false state existence and accepted it which is also analogues to “false negative”. Thus statistically, Type I error is when a null hypothesis is rejected when it is true, and Type II error is when a null hypothesis is accepted when it is false [1] [2].

Predictive value is used to measure the likelihood of true state. A Positive Predictive Value (PPV) is used to assess the proportion of samples which actually reflects the true state when in fact the diagnostic/experimental test indicates the presence of such state. A negative predictive value is also useful to determine the proportion of sample which is truly free when test indicates absence of a diseases or presence of a pathogen.

2.3. Mathematical Approach to Practical Measurements

From the principle of Neyman-Pearson Paradigm, Type I error occurs with a probability of α called “significance level” of the test and Type II error also occurs with a probability β , hence the probability of rejecting a null probability when it is indeed false is called the “Power” denoted as $1 - \beta$. Using equations to represents the concept, therefore, sensitivity is defined as [2]:

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (1)$$

Whereas specificity is also defined as:

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}} \quad (2)$$

False Positive rate (α)

$$\alpha = \frac{\text{False Positive}}{\text{False Positive} + \text{True Negative}} = 1 - \text{specificity} \quad (3)$$

False Negative rate (β)

$$\beta = \frac{\text{False Negative}}{\text{False Negative} + \text{True Positive}} = 1 - \text{sensitivity} \quad (4)$$

$$\text{Power} = 1 - \beta = \text{sensitivity} \quad (5)$$

Positive Predicted Value (PPV)

$$\text{PPV} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (6)$$

Negative Predicted Value (NPV)

$$\text{NPV} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Negative}} \quad (7)$$

Hence, the relation for sensitivity and specificity to hypothesis error is as shown in [Table 1](#).

3. Practical Application to Pathogen Enumeration Data

The study presents an attempt of interpretation of the use of a design methodology for the enumeration of *cryptosporidium* in the laboratory. In most cases of detection method, a recovery rate is calculated based on the detection strength of the method applied, and a corresponding control experiment is also carried out with a known spike of pathogen.

3.1. Material and Methods

The study was conducted on farms from four study sites, namely, Ahodwo, Chirepatre Estate, Twumduase and Boadi ([Figure 1](#)) all in the Kumasi Metropolis of the Ashanti Region of Ghana. Water samples were taken between April 2014 and January 2015. Permission to use the various sites for the study was granted by the waste

Table 1. Relationship of hypothesis error, sensitivity and specificity.

		State Conditions	
		Present	Absent
Test Results	Positive	Condition present + positive results = True positive	Condition absent + positive result = False positive (Type I error)
	Negative	Condition present + negative results = False negative (Type II error)	Condition absent + negative result = True negative

management department of the Kumasi Metropolitan Assembly, besides; farmers who own the farms where the study took place also granted us the permission to use their farms for the study. The field study does not involved endangered or protected species or protected area.

3.2. Water Sample Collection and Processing

All farms had different irrigation water sources; Farm (1) in Ahodwo, used irrigation water from a stream where wastewater from the Komfo Anokye Teaching Hospital (KATH) joins upstream without any proper treatment. Irrigation was performed using a pump as the water source is a little far from vegetable beds. Farm (2) in Chirepatre Estate, had two sources of irrigation water; hand dug well and stream which is joined upstream by effluent from waste stabilization pond. The areas supplying the dugout with water are mainly groundwater, private housing and run off from nearby green areas. Farm (3) in Twumduase, used two hand dug wells as the sole sources of irrigation. Farm (4) in Boadi, used a stream which is joined by various streams from surrounding communities (**Figure 1**). Collection of water samples was done once weekly from the farms over the period of study. Samples were taken from all water sources per farm thus Farms 2 and 3 which had two sources add up to give six sampling points in all. Twenty (20 L) litres of water sample was collected using two 10 L clean transparent graduated plastic containers from each sampling point and processed for *Cryptosporidium* spp. isolation as recommended [5]. Samples were taken from the water source 20 to 30 cm beneath the water surface. Water samples were transported to the Biochemistry Department of Kwame Nkrumah University of Science and Technology (KNUST) under optimal conditions and processed to obtain purified oocytes using immuno-magnetic separation (IMS).

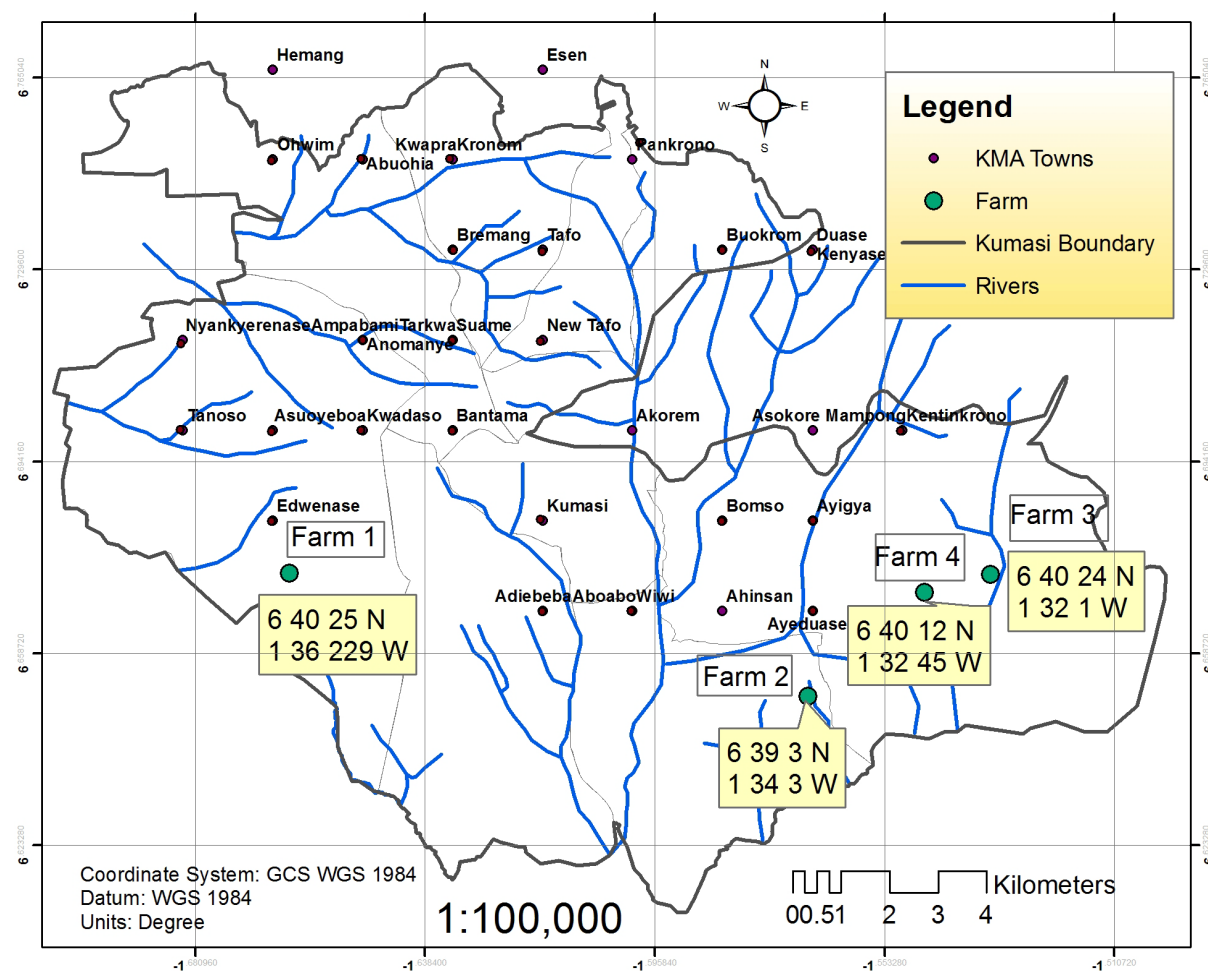


Figure 1. Farm sites where wastewater samples were collected in Kumasi, Ghana.

Purified oocysts were transported to the Department of Parasitology of the Noguchi Memorial Institute for Medical Research (NMIMR) for detection by Microscopy and Polymerase Chain Reaction (PCR).

3.3. Purification of *Cryptosporidium* spp. from Water Samples

The 10 L containers were left on the table in the laboratory for 48 hours, promoting sedimentation of *Cryptosporidium* spp. oocysts from top to bottom of the containers [6]. The supernatants were then removed by pump-suction system leaving approximately 0.75 L water in the containers. The remaining solutions were transferred to 3 L containers, followed by a 3×150 ml distilled water cleaning cycle (manual vortexing) of the 10 L containers. The 3 L containers were placed at a 30° angle and left on the table to sediment for another 48 hours. Consequently, the supernatants were removed, leaving approximately 90 ml in the containers. The remaining mixture was transferred to 50 ml tubes, followed by a 3×20 ml tap water clean-up cycle of the 3 L containers. The 50 ml tubes were centrifuged at 1583 g for 10 min and the supernatant removed leaving approximately 5 ml. Pellets were pooled in one of the 50 ml tubes followed by a 3×5 ml clean-up cycle with 0.01% Tween 20 in distilled water. The 50 ml tube containing the pellets and clean-up solution were centrifuged at 1585 g for 10 min and the supernatant removed leaving approximately 5 to 10 ml. The *Cryptosporidium* oocysts were purified using the Immune-Magnetic Separation (IMS) according to the manufacturer's protocol (Dynal Beads). After IMS, samples were aliquot appropriately for processing and detection of *Cryptosporidium* oocysts by microscopy and PCR.

3.4. Modified Ziehl Nelson Staining

After purification of *Cryptosporidium* from water samples, the resulting sediments were stained using the modified Ziehl Nelson staining as follows: After IMS, 100 μ l of each sample was smeared on glass slides and allowed to air dry followed by fixation in methanol for 3 min. The slides were then stained in carbol fuchsin for 15 to 20 min and then rinsed with tap water. The slides were decolourised in acid alcohol (1% HCl in methanol) for about 15 to 20 sec followed by thorough rinsing with water. Counterstaining was done with malachite green for 30 sec, rinsed and air dried. Slides were examined under x40 magnification using light microscope, to detect any oocysts.

3.5. Genomic DNA Extraction and Polymerase Chain Reaction (PCR)

Genomic DNA was extracted from processed wastewater using the Qiagen kit (QIAGEN Sciences, USA). The eluted genomic DNA was stored at -40°C until use. For molecular detection of *Cryptosporidium* (species/genotypes), PCR amplification of the HSP70 gene (325 bp) was done according to [7] with slight modifications. The PCR mix contained 12.5 μ l of Gotaq (Promega) and 12.5 pmol of each primer in a total reaction volume of 25 μ l. The PCR mix included 1 to 5 ml of purified DNA as template for primary steps and 1 μ l of primary PCR product for secondary steps. The PCR was carried out in a Takara thermocycler with an initial hot start (95°C for 3 min) and a final extension (72°C for 10 min). For the nested PCR amplification of the HSP70 gene, 30 cycles (94°C for 30 sec, 58°C for 20 sec, 72°C for 30 sec) cycling conditions were set using the HSP4 primers for the primary step and 45 cycles (94°C for 25 sec, 58°C for 18 sec, 72°C for 25 sec) using HSP3m primers in the secondary step. Positive and negative controls were included in every reaction.

4. Analysis and Results

Maximum likelihood method was used to estimate specificity and the sensitivity of Microscopic and PCR methods. 24 water samples were stained with cryptosporidium to represent the presence of cryptosporidium samples in water whiles 20 water samples free of *Cryptosporidium* were also used as a control to evaluate the processes of both microscopic and PCR.

A total of 9 samples were tested positive out of the 24 positive samples using the Microscopic method, and 15 tested negative. For the control, none of the samples of *Cryptosporidium* tested positive to indicate a false positive, thus all the 20 water samples fails to response to positive presence of *Cryptosporidium* (true negative, **Table 2**). On the part of PCR, 8 samples recorded for true positive, indicating positive results for water samples with *Cryptosporidium* whiles all the 20 water samples shows true negative results for the control process (**Table 3**). By Comparison of Microscopic and PCR results for sensitivity and specificity. The study recorded 37.5%

Table 2. Results for microscopic.

		State Conditions		Total
		Present	Absent	
Test Results	Positive	9	0	9
	Negative	15	20	35
	Total	24	20	44

Table 3. Results for PCR.

		State Condition of Cryptosporidium		Total
		Present	Absent	
Test Results	Positive	8	0	8
	Negative	16	20	36
	Total	24	20	44

sensitivity for microscopic and 33.3% for PCR and also 100.00% specificity for both microscopic and PCR (Table 4).

5. Discussions

In statistical hypothesis testing, false positive rate is a function of specificity and prevalence rate for testing experimental units. Specifically, specificity and sensitivity relate to correct decisions in statistical hypothesis, while false positive and false negative rates lead to type I and type II error analysis. In general, a “true positive” test results in experimental work leads to a significant result affecting the degree of failing to accept a null hypothesis while a “true negative” test results is a vice versa. It is noteworthy that, a high “true positive” leads to a high probability of detecting a false null hypothesis and measures the sensitivity; on the other hand, a high value of “true negative” has a high probability accepting a true null hypothesis and measures the specificity of the study. It should be noted that, in experimental settings, a low value for false positives leads to a low probability of type I error while a low value of false negatives leads to low probability of committing type II error.

In this study, the measure of sensitivity was low for both Microscopic and PCR (thus 37.5% and 33.3% respectively) which lowers the probability of detecting a “false null hypothesis” for the presence of *cryptosporidium* in the water samples using either Microscopic or PCR. Nevertheless, both procedures for *cryptosporidium* detection had a higher “true negatives” increasing its probability of failing to reject a “true null hypothesis” with specificity of 1.00 for both Microscopic and PCR. Results values of experiments can skew to any direction with respect to sensitivity and specificity, some process can have both higher sensitivity and specificity at the same time. In the staining method in *cryptosporidium parvum* diagnosis in diarrheic sample of patients, Reference [8] recorded 83.7% sensitivity and 98.9% specificity for staining method which represents a high probability for rejecting and accepting a true null hypothesis in both case. On the other hand, recording a higher sensitivity and lower specificity is also possible, Reference [9] evaluation of four different methods (Ziehl neelsen staining, safranin methylen blue staining, Ag detection and Nested-PCR) for *Cryptosporidium* detection had a result of 100% sensitivity for PCR and 41.5% sensitivity for MZN staining, with a specificity results of 100% for both PCR and MZN, however, a lower sensitivity and higher specificity are in both cases possible as seen in some studies [10]. Reference [11] compared the sensitivity and specificity of a modified Ziehl-Neelsen (modified-ZN) staining method for acid-fast bacilli (AFB) with that of the standard Ziehl-Neelsen (standard-ZN) staining method, and recorded a sensitivity 72% (101 of 140) and 84% (117 of 140); of the modified-ZN staining method and standard-ZN staining method the modified-ZN method missed 21% of cases detected by the standard-ZN method. Similar results [12]-[16] were also recorded. This indicates the non-directionless of specificity and sensitivity indicating they are independent in experimental detection methods. Above all these, the most difficult question to answer is to quantify the term “answering the wrong question” this is referred to as type III error. This error is as a result of ignoring the most important variable affecting a process performance in the experimental design. Another form is when the result is much improved but the resulting loss of purity prevents successful extraction of target ingredients, these are difficult to deal with and cannot be improved with increasing sample size.

Table 4. Sensitivity and specificity results.

	Microscopic	PCR
Sensitivity	0.375	0.333
Specificity	1.000	1.000
False Positive Rate	0.000	0.000
False Negative Rate	0.625	0.667
Positive Predictive Value	1.000	1.000
Negative Predictive Value	0.571	0.556

6. Conclusion

Measuring the sensitivity and specificity of experimental procedure is never a general rule for a general statement with respect to prevalence of a procedure. Universal accepted procedures for detection of pathogens sensitivity and specificity analysis cannot be generalized since such procedures are affected by both environmental and human influences. Sensitivity, specificity, false negative rate and false positive rate are essential to measure the probability of accepting a true null hypothesis, or rejection of a false null hypothesis as well as quantifying the likelihood of making either type I or type II error in statistical hypothesis with respect to a localized experiment. As a general thumb of rule, a higher sensitivity, specificity or a lower false negative rate, false positive rate of a particular experiment with a specific detection method does not necessarily indicate an overall efficiency of the pathogen detection procedure unless it's related to a gold standard. In summary, sensitivity and specificity are properties for indicating a degree of reliability of diagnostic/experimental test and do not predict the predictive value. Sensitivity and specificity are merely properties of a test and should not be used to make general statement as findings.

References

- [1] Bennet, M.B. (1972) On Comparisons of Sensitivity, Specificity and Predictive Value of a Number of Diagnostic Procedures. *Biometrics*, **28**, 793-800. <http://dx.doi.org/10.2307/2528763>
- [2] Gaddis, G.M. and Gaddis, M.L. (1990) An Introduction to Biostatistics Part 3: Sensitivity, Specificity, Predictive Value and Hypothesis Testing. *Annals of Emergency Medicine*, **19**, 591-597. [http://dx.doi.org/10.1016/S0196-0644\(05\)82198-5](http://dx.doi.org/10.1016/S0196-0644(05)82198-5)
- [3] Sharma, D., Yadav, U.B. and Sharma, P. (2009) The Concept of Sensitivity and Specificity in Relation to Two Types of Errors and Its Application in Medical Research. *Journal of Reliability and Statistical Studies*, **2**, 53-58.
- [4] Kendall, M., Stuart, A., Ord, K.J. and Arnold, S. (1999) Kendall's Advanced Theory of Statistics: Volume 2A Classical Inference and the Linear Model (Kendall's Library of Statistics). 6th Edition, Hodder Arnold Publication, London.
- [5] Duhain, G. (2011) Occurrence of *Cryptosporidium* spp in South African Irrigation Waters and Survival of *Cryptosporidium parvum* during Vegetable Processing. University of Pretoria, South Africa.
- [6] Medema, G., Bahar, M. and Schets, F. (1997) Survival of *Cryptosporidium parvum* *Escherichia coli*, *Faecal enterococci* and *Clostridium perfringens* in River Water: Influence of Temperature and Autochthonous Microorganisms. *Water Science & Technology*, **35**, 249-252. [http://dx.doi.org/10.1016/S0273-1223\(97\)00267-9](http://dx.doi.org/10.1016/S0273-1223(97)00267-9)
- [7] Langkjaer, R.B., Vigre, H., Enemark, H.L. and Maddox-Hyttel, C. (2007) Molecular and Phylogenetic Characterization of *Cryptosporidium* and *Giardia* from Pigs and Cattle in Denmark. *Parasitology*, **134**, 339-350. <http://dx.doi.org/10.1017/S0031182006001533>
- [8] Morgan, U.M., Pallant, L., Dwyer, B.W., Forbes, D.A., Rich, G. and Thompson, R.C. (1998) Comparison of PCR and Microscopy for Detection of *Cryptosporidium parvum* in Human Fecal Specimens: Clinical Trial. *Journal of Clinical Microbiology*, **36**, 995-998.
- [9] Kaushik, K., Khurana, S., Wanchu, A. and Malla, N. (2008) Evaluation of Staining Techniques, Antigen Detection and Nested PCR for the Diagnosis of Cryptosporidiosis in HIV Seropositive and Seronegative Patients. *Acta Tropica*, **107**, 1-7. <http://dx.doi.org/10.1016/j.actatropica.2008.02.007>
- [10] Omoruyi, B.E., Nwodo, U.U., Udem, C.S. and Okonkwo, F.O. (2014) Comparative Diagnostic Techniques for Cryptosporidium Infection. *Molecules*, **19**, 2674-2683. <http://dx.doi.org/10.3390/molecules19022674>
- [11] Selvakumar, N., Rahman, F., Rajasekaran, S., Narayanan, P.R. and Frieden, T.R. (2002) Inefficiency of 0.3% Carbol Fuchsin in Ziehl-Neelsen Staining for Detecting Acid-Fast Bacilli. *Journal of Clinical Microbiology*, **40**, 3041-3043.

<http://dx.doi.org/10.1128/JCM.40.8.3041-3043.2002>

- [12] Chalmers, R.M., Campbell, B.M., Crouch, N., Charlett, A. and Davies, A.P. (2011) Comparison of Diagnostic Sensitivity and Specificity of Seven *Cryptosporidium* Assays Used in the UK. *Journal of Medical Microbiology*, **60**, 1598-1604. <http://dx.doi.org/10.1099/jmm.0.034181-0>
- [13] Shimelis, T. and Tadesse, E. (2014) Performance Evaluation of Point-of-Care Test for Detection of *Cryptosporidium* Stool Antigen in Children and HIV Infected Adults. *Parasit. Vectors*, **7**, 227. <http://dx.doi.org/10.1186/1756-3305-7-227>
- [14] Marques, F.R., Cardoso, L.V., Cavasini, C.E., De Almeida, M.C., Bassi, N.A., De Almeida, M.T.G., Rossit, A.R.B. and Machado, R.L.D. (2005) Performance of an Immunoenzymatic Assay for *Cryptosporidium* Diagnosis of Fecal Samples. *Brazilian Journal of Infectious Diseases*, **9**, 3-5. <http://dx.doi.org/10.1590/S1413-86702005000100002>
- [15] Tahvildar-Biderouni, F. and Salehi, N. (2014) Detection of *Cryptosporidium* Infection by Modified Ziehl-Neelsen and PCR Methods in Children with Diarrheal Samples in Pediatric Hospitals in Tehran. *Gastroenterology and Hepatology from Bed to Bench*, **7**, 125-130.
- [16] Shea, Y.R., Davis, J.L., Huang, L., Kovacs, J.A., Masur, H., Mulindwa, F., Opus, S., Chow, Y. and Murray, P.R. (2009) High Sensitivity and Specificity of Acid-Fast Microscopy for Diagnosis of Pulmonary Tuberculosis in an African Population with a High Prevalence of Human Immunodeficiency Virus. *Journal of Clinical Microbiology*, **47**, 1553-1555. <http://dx.doi.org/10.1128/JCM.00348-09>