# Misclassification in defining and diagnosing microcephaly

**Mariah M. Kalmin**[1], **Emily Gower**[1], **Elizabeth Stringer**[2], **Natalie M. Bowman**[3], **Elizabeth T. Rogawski**[4], **Daniel Westreich**[1]

[1]Department of Epidemiology, UNC-Chapel Hill, Chapel Hill, NC

[2]Department of Obstetrics and Gynecology, UNC-Chapel Hill, Chapel Hill, NC

[3]Department of Medicine, UNC-Chapel Hill, Chapel Hill, NC

[4]Department of Public Health Sciences, University of Virginia, Charlottesville, VA

## Abstract

**Background:** Several health agencies define microcephaly for surveillance purposes using a single criterion, a percentile or Z-score cutoff for newborn head circumference. This definition, however, conflicts with the reported prevalence of microcephaly even in populations with endemic Zika virus.

**Objective:** We explored possible reasons for this conflict, hypothesizing that the definition of microcephaly used in some studies may be incompletely described, lacking the additional clinical criteria that clinicians use to make a formal diagnosis. We also explored the potential for misclassification that can result from differences in these definitions, especially when applying a percentile cutoff definition in the presence of the much lower observed prevalence estimates that we believe to be valid.

**Methods:** We conducted simulations under a theoretical bimodal distribution of head circumference. For different definitions of microcephaly, we calculated the sensitivity and specificity using varying cutoffs of head circumference. We then calculated and plotted the positive predictive value for each of these definitions by prevalence of microcephaly.

**Results:** Simple simulations suggest that if the true prevalence of microcephaly is approximately what is reported in peer-reviewed literature, then relying on cutoff-based definitions may lead to very poor positive predictive value under realistic conditions.

**Conclusions:** While a simple head circumference criterion may be used in practice as a screening or surveillance tool, the definition lacks clarification as to what constitutes true pathological microcephaly and may lead to confusion about the true prevalence of microcephaly in Zika-endemic areas, as well as bias in etiologic studies.

Correspondence Mariah M. Kalmin, Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, mkalmin@email.unc.edu. Alternate Correspondence Daniel Westreich (ibid.), westreic@email.unc.edu.

Conflict of interest None declared.

## Background

Since 2015 microcephaly has received increased awareness because of its association with outbreaks of Zika virus throughout most of Central and South America.[1–4] Microcephaly occurs when a child's head is smaller than typical and is usually accompanied by impaired brain growth and development.[5–7] Canonical definitions of microcephaly include a newborn or fetal head circumference below the third percentile or, similarly, a head circumference at least two standard deviations (SDs) below the mean from a reference population, for age and gender.[5, 8] Both definitions of microcephaly have been used by the Centers for Disease Control and Prevention (CDC) (although the CDC alludes in some places to the need for additional, confirmatory evidence)[5] and American Academy of Pediatrics (AAP), as well as throughout the peer-reviewed literature.[5, 8] The Society for Maternal-Fetal Medicine (SMFM) has recommended that isolated fetal microcephaly should be defined as a fetal head circumference (as measured through ultrasound) at least three SDs below the mean for gestational age.[9]

Based on the definition that relies solely on a third percentile or two SD cutoff, we should expect approximately 2–3% of newborns to have microcephaly.[5, 8] Additionally, if we are studying a population specifically because we suspect an excess number of cases (as in the setting of Zika epidemics), we would expect to observe an even greater prevalence using such a definition. However, several recent surveillance studies from South America that cited these definitions have reported a far lower prevalence of microcephaly, with estimates varying between 1 and 12 in 10,000 children (from 0.01% to 0.12%).[10–13] Of these four studies, two specify that additional clinical, laboratory, or radiological criteria are required for case confirmation though these criteria are not specified.[11, 12] In the two remaining studies, there is no further explanation of the criteria used to diagnose microcephaly beyond the head circumference cutoff.[10, 13] Surveillance studies conducted throughout the US during a similar time period have emphasized the need for additional neuroimaging and clinical diagnosis, but the exact types of clinical evaluation were not always specified or consistent across these studies.[14–16]

This gap between the expected 2–3% prevalence, derived from the cutoff-based definition, and the reported prevalence of 0.01 to 0.12% in some studies could be due to several factors. For example, this discrepancy could result from a larger mean head circumference in the studied populations compared to the reference populations, an increased level of protective factors that reduce the risk of microcephaly in the studied populations, or systematic underreporting of microcephaly. Additionally, measuring newborn head circumference in the typical hospital setting may be error prone due to difficulty in obtaining accurate measurements and changes in head shape following birth.[2] And using brain imaging (sometimes repeatedly over time) to confirm a diagnosis may not be feasible during large scale Zika epidemics or in resource-limited settings. It is notable, as well, that the reported

prevalence from these studies is more consistent with the SMFM definition (head circumference cutoff for fetal microcephaly three SDs below the mean, about 0.135% of a normal distribution).[9]

We hypothesize, however, that the prevalence estimates reported in studies are valid and that the differences between the expected prevalence derived from the cutoff-based definition and the reported prevalence estimates are due to the use of additional, sometimes-unreported diagnostic criteria beyond newborn head circumference. As noted, in some studies, classification of an infant with microcephaly has been reported to include diagnosis of brain abnormalities (intracranial calcifications, cerebral atrophy, etc.)[15] in conjunction with small head circumference; however, these diagnostic criteria are sometimes left unspecified and remain unclear.

The confusion in the extant literature gives rise to the possibility that some investigators may use (perhaps inadvertently) a cutoff-based definition of microcephaly in research. While such sensitive definitions may be appropriate for use in large-scale surveillance, they may lead us astray in research that requires more specific identification of microcephaly cases. For example, misclassification of pathological microcephaly equally among Zika-exposed and Zika-unexposed infants (that is, nondifferential outcome misclassification) could lead to biased estimates of the causal effect of Zika virus on pathological microcephaly, likely towards the null if the misclassification is non-differential with respect to exposure status. In the remainder of this report, we explore the scale of misclassification bias that may be introduced through using a more sensitive cutoff-based definition, if we assume that the more specific, clinical definitions represent true pathological microcephaly prevalence.

## Methods

Suppose that the distribution of newborn head circumference takes on a bimodal distribution (Figure 1), of which a proportion of the population has microcephaly (left mode of the distribution, centered over −4) and the remainder of the population does not have microcephaly and is considered healthy (right mode, centered over 0).

To illustrate how the current definition of microcephaly performs under such a bimodal distribution of head circumference, we constructed several scenarios. For the true pathological distribution, we set the mean at different Z-scores (−3, −4, −5), with cutoffs for defining pathological microcephaly at Z-scores of −2, −3, and −4; the mean of the pathological distribution was constrained to be always less than the cutoff (Table 1). We then calculated the sensitivity and specificity for each of these six scenarios. To understand the implications of varying sensitivity and specificity, we plotted the positive predictive value (PPV) by different estimates of pathological microcephaly prevalence in the entire population. We set the prevalence at 0.01%, 0.1%, 1%, and 10% and then calculated the PPV (Figure 2).

## Results

Generally, as the distance between the *pathological mean* (the mean of the true pathological distribution) and the testing cutoff grows, more true-positive cases of pathological

microcephaly are captured allowing the sensitivity to increase. The greatest sensitivity is achieved in scenario D, where the pathological mean is set to a Z-score of −5 and the Z-score cutoff for testing for pathology is −2 (three SDs between the mean and cutoff). The sensitivity is lowest in scenarios A, C, and F where there is only one SD between the pathological mean and testing cutoff. On the other hand, the specificity increases as the cutoff for pathological microcephaly gets farther from 0 (the *healthy mean*, the mean of the truly healthy individuals) since fewer false positive cases are captured.

The highest PPV of 99.97% was achieved at a prevalence of 10% in scenario F, since the specificity is 99.99997% (very few false positives) (Figure 2F). In scenarios C and E, where the specificity is 99.9%, we see PPVs of at least 90% starting at 1% prevalence, but in scenarios A, B, and D, the highest PPV achieved is capped between 80–83%, and only when the prevalence of pathological microcephaly is highest at 10%. Notably, under what we believe might be a realistic scenario (e.g., true prevalence 0.01% to 0.1%; pathological mean = −4; cutoff = −2) PPV is extremely low at 0.42% to 4.1% (Figure 2B).

## Comment

The number of cases of microcephaly reported in recent publications is lower than we would expect given formal cutoff definitions for microcephaly. This suggests the possibility of misclassification in either the formal cutoff definitions (which may have high sensitivity but low specificity) or the reported cases (which may have high specificity but low sensitivity). Simulations demonstrate that if the true prevalence of microcephaly is approximately what is being reported in peer-reviewed literature,[10–13] then relying on the formal cutoff definitions may lead to very poor positive predictive value under realistic conditions.

While the assumptions of normality (or bimodality) in our illustrations may be violated in observed populations, the exact numbers reported in the Figures and Table are meant to illustrate probable misclassification using the more sensitive, but less specific cutoff-based definition of microcephaly; similar principles would apply with non-normal distributions. We do not discount the importance of growth metrics, such as head circumference, to indicate health status (especially in low-resource settings), due to the relative ease with which they are measured. We suggest, instead, further transparency in the reporting of methods used for defining microcephaly, and highlight the need to further research around this methodological issue.

## Conclusions

In conclusion, clinicians and researchers should recognize that the current official cutoff-based definitions of microcephaly – while potentially useful in surveillance contexts – are difficult to resolve against reported prevalence figures in the literature. If this contradiction is because of the use of additional diagnostic criteria in clinical practice, then documentation of these additional criteria in the reported definition of microcephaly will be an important contribution to our understanding of microcephaly and reproducibility across studies. On the contrary, if the discrepancy is due to under-reporting and delays in diagnosis, then surveillance efforts should be assessed and improved in order to capture the full scope of

disease. Furthermore, it is important to recognize that the disconnect between the definition of microcephaly and observed prevalence estimates has the potential to lead to biased estimates of the association between Zika virus (or other possible causes) and microcephaly. We suggest further methodological work in this field to elucidate the most valid definitions of microcephaly that minimize misclassification and subsequent biases.

## Acknowledgements

## References

1. Ribeiro IG, Andrade MR, Silva JM, Silva ZM, Costa MAO, Vieira M, et al. Microcephaly in Piaui, Brazil: descriptive study during the Zika virus epidemic, 2015–2016. Epidemiol Serv Saude. 2018; 27:e20163692.

2. Nunes ML, Carlini CR, Marinowic D, Neto FK, Fiori HH, Scotta MC, et al. Microcephaly and Zika virus: a clinical and epidemiological analysis of the current outbreak in Brazil. J Pediatr (Rio J). 2016; 92:230–240. [PubMed: 27049675]

3. Coelho AVC, Crovella S. Microcephaly Prevalence in Infants Born to Zika Virus-Infected Women: A Systematic Review and Meta-Analysis. Int J Mol Sci 2017; 18.

4. de Araújo TVB, Ximenes RAA, Miranda-Filho DB, Souza WV, Montarroyos UR, de Melo APL, et al. Association between microcephaly, Zika virus infection, and other risk factors in Brazil: final report of a case-control study. Lancet Infect Dis. 2018; 18:328–336. [PubMed: 29242091]

5. Centers for Disease Control and Prevention. Facts about Microcephaly. 2016 [February 9, 2018]; Available from: https://www.cdc.gov/ncbddd/birthdefects/microcephaly.html.

6. World Health Organization. Microcephaly Fact Sheet. 2016 [February 9, 2018]; Available from: http://www.who.int/mediacentre/factsheets/microcephaly/en/.

7. National Institute of Neurological Disorders and Stroke. Microcephaly Information Page. 2017 [February 9, 2018]; Available from: https://www.ninds.nih.gov/Disorders/All-Disorders/Microcephaly-Information-Page.

8. McInerny TK, Adam HM, Campbell DE, DeWitt TG, Foy JM, Kamat DM, eds. American Academy of Pediatrics Textbook of Pediatric Care. 2nd ed Elk Grove Village, IL: American Academy of Pediatrics; 2017.

9. Society for Maternal-Fetal Medicine Publications C. Ultrasound screening for fetal microcephaly following Zika virus exposure. Am J Obstet Gynecol 2016; 214:B2–4. [PubMed: 26901275]

10. DeSilva M, Munoz FM, Sell E, Marshall H, Tse Kawai A, Kachikis A, et al. Congenital microcephaly: Case definition & guidelines for data collection, analysis, and presentation of safety data after maternal immunisation. Vaccine. 2017; 35:6472–6482. [PubMed: 29150052]

11. Almeida I, Ramos CV, Rodrigues DC, Sousa AC, Nascimento M, Silva M, et al. Clinical and epidemiological aspects of microcephaly in the state of Piaui, northeastern Brazil, 2015–2016. J Pediatr (Rio J). 2018.

12. Magalhaes-Barbosa MC, Prata-Barbosa A, Robaina JR, Raymundo CE, Lima-Setta F, Cunha AJ. Trends of the microcephaly and Zika virus outbreak in Brazil, January-July 2016. Travel Med Infect Dis. 2016; 14:458–463. [PubMed: 27702683]

13. Tellechea AL, Luppo V, Morales MA, Groisman B, Baricalla A, Fabbri C, et al. Surveillance of microcephaly and selected brain anomalies in Argentina: Relationship with Zika virus and other congenital infections. Birth Defects Res. 2018.

14. Cragan JD, Isenburg JL, Parker SE, Alverson CJ, Meyer RE, Stallings EB, et al. Population-based microcephaly surveillance in the United States, 2009 to 2013: An analysis of potential sources of variation. Birth Defects Res A Clin Mol Teratol. 2016; 106:972–982. [PubMed: 27891783]

15. Delaney A, Mai C, Smoots A, Cragan J, Ellington S, Langlois P, et al. Population-Based Surveillance of Birth Defects Potentially Related to Zika Virus Infection - 15 States and U.S. Territories, 2016. MMWR Morb Mortal Wkly Rep. 2018; 67:91–96. [PubMed: 29370151]

16. Rice ME, Galang RR, Roth NM, Ellington SR, Moore CA, Valencia-Prado M, et al. Vital Signs: Zika-Associated Birth Defects and Neurodevelopmental Abnormalities Possibly Associated with Congenital Zika Virus Infection - U.S. Territories and Freely Associated States, 2018. MMWR Morb Mortal Wkly Rep. 2018; 67:858–867. [PubMed: 30091967]

**Social media quote**

The number of microcephaly cases reported in recent publications is lower than expected given formal population percentile cutoff definitions for microcephaly, suggesting possible misclassification in either the formal cutoff definitions or the reported cases.

**Synopsis**

**Study Question:**

What are the implications of conducting epidemiologic research using the current definition of microcephaly that relies on a cutoff of newborn or fetal head circumference?

**What's already known:**

Microcephaly is a rather rare condition with reported prevalence estimates varying between 1 to 12 per 10,000 children even though it is defined as a newborn or fetal head circumference less than the third percentile in the population.

**What this study adds:**

This study addresses the discrepancy between the definition of microcephaly based on newborn or fetal head circumference and the reported estimates of microcephaly prevalence.
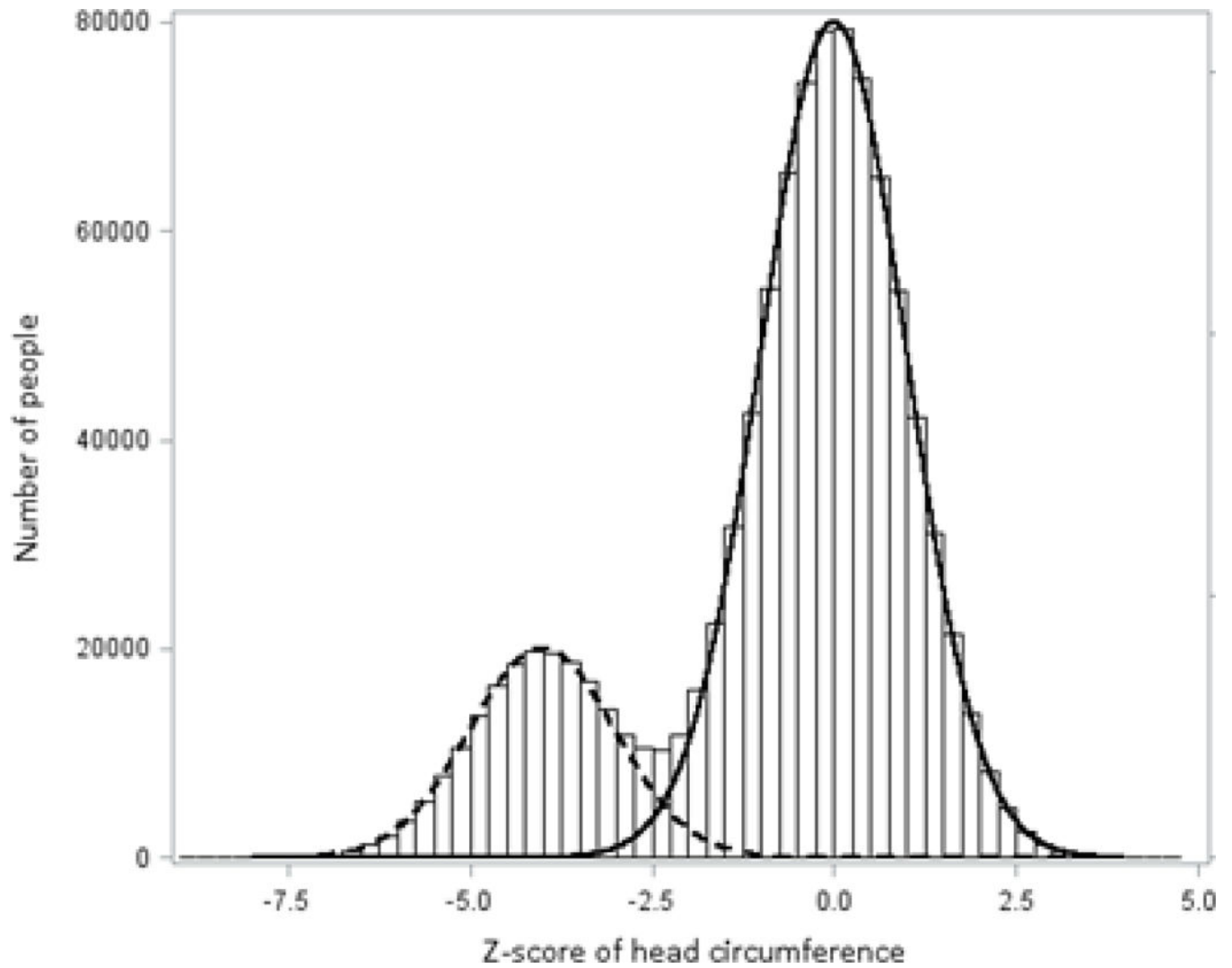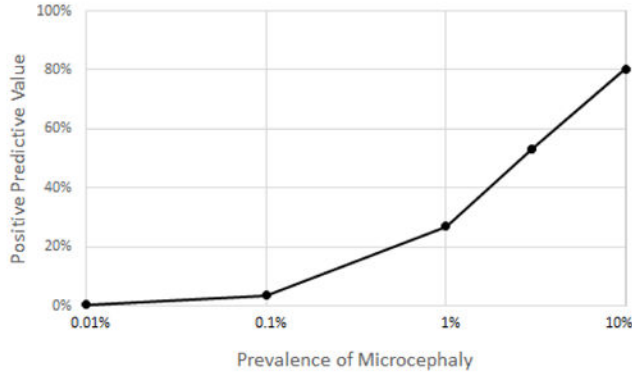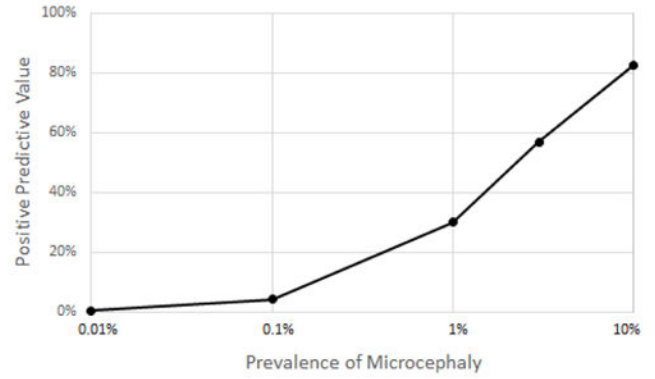
**Figure 1. Example of a bimodal distribution with superimposed normal curves.**
The histogram is the bimodal distribution with 20% at mean –4, SD 1 and 80% at mean 0,
SD 1. The two curves are both normal curves with mean (SD) as follows: dashed line, –4
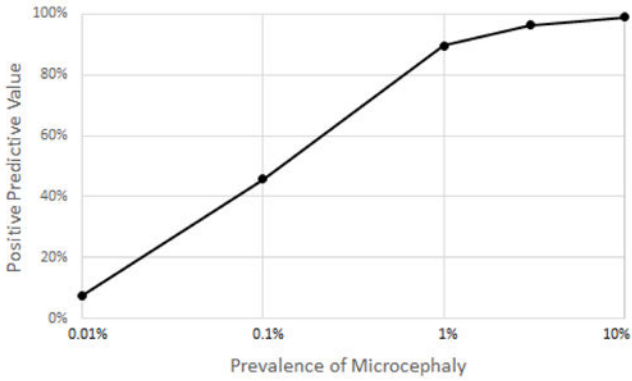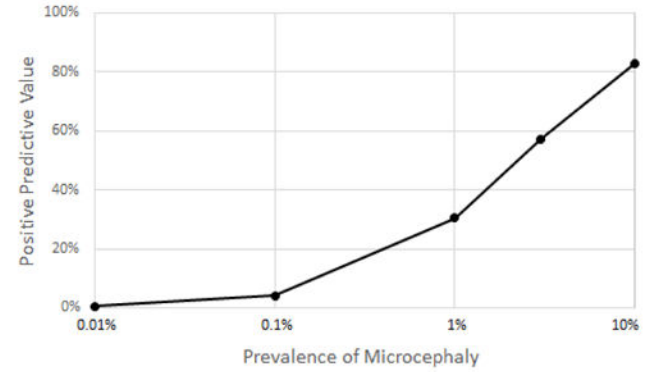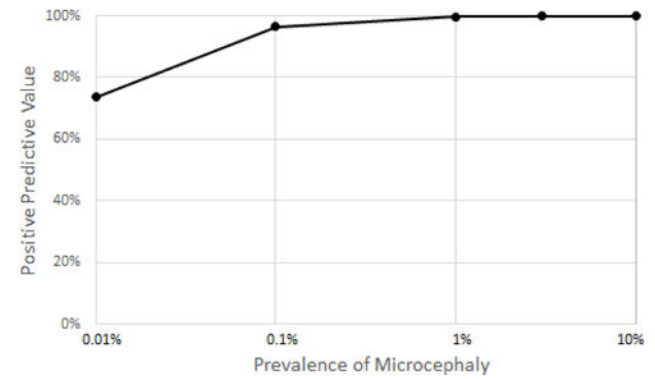(1); solid line, 0 (1).
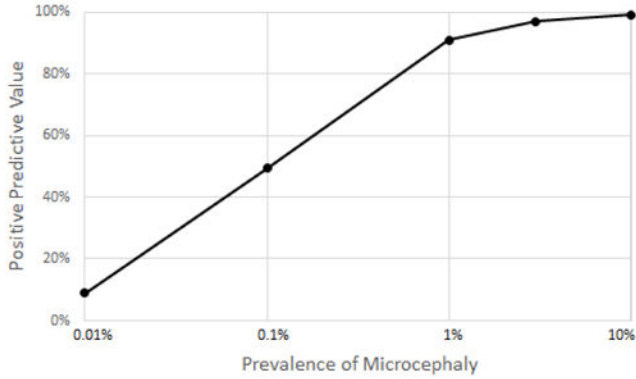
**Figure 2. Positive predictive value and prevalence estimates with varying sensitivity (Se) and specificity (Sp).**

Note that the x-axis on all panels is shown in log-scale. Scenarios in each panel correspond to rows labeled with the same letters in Table 1.

**Table 1.**

**Sensitivity and specificity for different definitions of pathological head circumference.**

Each row of the table estimates sensitivity and specificity assuming two normal distributions: one for non-pathological newborns (not shown mean 0 for all rows; SD 1), and one for true pathology (mean as shown by row; SD 1). Third column of table shows Z-score cutoff at which we define pathology. Thus, in scenario A, we report sensitivity and specificity when (i) head circumference among truly microcephalic newborns has a normal distribution with mean −3, SD 1; and (ii) if we define pathology as "any head circumference below a Z-score of −2." Note that sensitivity changes with the relationship between true pathology and pathological cutoff, but (since non-pathological heads have the same distribution in every case) specificity changes with pathological cutoff *only*.

| Scenario | True pathology (mean Z-score) | Testing for pathology (Z-Score cutoff) | Diagnostic Test Characteristics | |
|----------|-------------------------------|------------------------------------------|-----------------------------------|-----------|
| | | | Sensitivity | Specificity |
| A | −3 | −2 | 84.1% | 97.7% |
| B | −4 | −2 | 97.7% | 97.7% |
| C | −4 | −3 | 84.1% | 99.9% |
| D | −5 | −2 | 99.9% | 97.7% |
| E | −5 | −3 | 97.7% | 99.9% |
| F | −5 | −4 | 84.1% | 100.0% [a] |

[a] True value is 99.99997%.