

CHAPTER 6.

CONCLUSIONS AND FUTURE IMPLICATIONS

Except otherwise noted, this work is made available under a Creative Commons Attribution License. <http://creativecommons.org/licenses/by/4.0>

Denno, DM, VanBuskirk, KM, Nelson, ZC, Musser, CA, Tarr, PI. (2016). Environmental enteric dysfunction: Advancing current knowledge. St. Louis, MO: Washington University Libraries. <http://dx.doi.org/10.7936/K7WQ0228>

Chapter 6. Conclusions and Future Implications

6.1 Summary of Findings

The 77 references from the years 2000 to 2010 provide a robust body of knowledge on biomarkers and diagnostic tests that could be useful in the quest to identify EED pathology, especially actionable EED, the entity most in need of identification to mitigate its long-term and life-threatening consequences.

Markers of permeability were the most frequently utilized biomarkers in the studies that we reviewed and the urinary L:M test was the most common of those measures. Lactoferrin, as well as other stool markers of intestinal inflammation, were also frequently assessed. Marker results varied greatly in terms of indicating the presence of intestinal dysfunction and degree of dysfunction when positive. This is not unexpected given the diverse populations that were studied, including different settings and subject presentations (e.g., asymptomatic children as well as those with various symptoms). In 20 of the 25 studies using the L:M test, the permeability ratio was noted to be elevated. As mentioned in the L:M section in Chapter 5, interpretation and comparison across studies in this review were hampered by variations in the way data were reported.

6.2 Future Biomarker and Diagnostics Research

This analysis was a comprehensive attempt to identify the most promising biomarkers or diagnostic tests on which to base future efforts and analyses of EED. Our review found many different biomarkers, across a wide range of diagnostic categories that have been deployed to

assess enteric function in children in resource-poor settings. However, regrettably, the extant literature does not permit us to recommend a specific technology or biomarker that can be used as a standard for the diagnosis of EED. There are many reasons for this difficulty:

1. The “tests” and the studies in which they were used were not developed with a biomarker concept in mind. The assays were performed chiefly to corroborate processes and were related to various host characteristics, but the relation to gut dysfunction was inferred without validation. While the preponderance of tests reported plausibly relate to gut dysfunction, their association with this particular pathophysiology was not proven.
2. The disorder of interest is poorly defined. It is clear that there is considerable and convergent evidence of abnormal digestive systems in the population of interest (children in low-income settings), and that there is an outcome (i.e., stunting) that is conceivably related to dysfunction of the gut, but the link has not been made conclusively.
3. Histopathological evidence of gut inflammation, which has been considered to prove the presence of EED, is an elusive gold standard (see discussion below on small bowel biopsies). Only two studies related histopathology to specific extra-intestinal markers. The L:M was found to correlate with intestinal inflammatory cell infiltration [111], while D-xylose was not associated with degree of villous atrophy [136]. Moreover, there is no evidence that histopathology is related to the ultimate outcome of interest, namely stunting, or other consequential host injury such as vaccine failure or increased susceptibility to infections.
4. The data presented in the publications that we analyzed often did not reflect high-grade subject reporting. For example, we often encountered data reported only in

terms of p-values and lacking reporting of effect sizes. In the field of randomized controlled trials, adherence to Consolidated Standards of Reporting Trials (CONSORT) Statement guidelines [309, 310] has improved the quality of clinical trials [311-313]. The effect of the non-systematic reporting of data in the biomarker field is that populations are not well described, methodology is not always discernible, and statistical tests are not completely described. The data that result might either obscure, or overstate, the value of the tests being assessed. We applaud efforts to standardize the quality of reporting of tests and biomarkers [314].

5. More data were certainly available than were reported. For example, there is value to the separate and simultaneous reporting of the lactulose and mannitol clearances, but such values were often not provided. Some studies reported a biomarker as significantly associated with an outcome of interest or correlated with another marker. However, often they did not mention whether other outcomes or biomarkers examined in the study were related to the biomarker of interest, leaving the reader unsure of whether these relationships were assessed at all or if only select (e.g., positive) findings were reported. Correlation analyses between the tests employed in the study were rarely reported. The publications that we reviewed did not report utilization of supplementary data deposition resources (e.g., supplementary or web-linked tables). However, such on-line registries for meta-data have only recently been made widely available by journals, better enabling capacity for all data to be fully utilized to answer questions that may not have been primary to the published study.
6. The biomarker studies that we analyzed, in the main, were often research project-driven. They were rarely scrutinized in the context of modern test-development standards. Many potential aspects critical to the clinical usefulness of a test, and to

the interpretation of the reported data, were not addressed. These include matters pertaining to the following: specimen transport, physical condition of the specimen (e.g., water content of stool), specimen processing prior to assay performance, temperature sensitivity, dynamic range of the test result, inhibitors (or molecules that could cause false positives) in the analyte, durability of the assays, inter-laboratory variability, biologic perturbations in specimens that often contain bacteria (stool and possibly urine), and assessments of the distribution of values in normal groups. An additional concern is that the target biomarker is presumed to be stable in the body fluid assayed, but this is rarely documented. We would encourage the broader use of spiking experiments, whereby the target molecule is added to the specimen of interest, and, ideally, recovered quantitatively. These study design, subject characterization, laboratory and analyte specification, and analysis and reporting considerations should be well addressed in future biomarker development work.

The multitude of test methods and procedures utilized, the disparate nature of the studies, and the limited testing for correlation with outcomes of interest (e.g., clinical signs of EED, histopathology or other markers of EED, or clinical outcomes of EED, such as stunting,) restrict our ability to declare any markers as clear lead candidates that warrant major or exclusive investment in the next phase of EED biomarker research.

It is important to note that our analysis was focused only on literature produced during the last eleven years. We acknowledge that it is possible that additional information could be extracted from studies published during prior years. In fact, we identified more references that were potentially relevant to our review question from the period 1990-1999 than for the period 2000-2010. However, a review of 10 randomly selected articles published between 1990-1999 and relevant to our review question (Appendix 3) did not add significant information compared to the 77 more recent publications included in the systematic review. Also, the aforementioned

structural issues in the reporting of the data in the studies from the 11-year period used in this systematic review (e.g., small sample sizes, study design and reporting issues) were also evident in the earlier studies. Moreover, even though EED research was more prevalent in the 1970s and 1980s, secular trends in nutritional, socioeconomic, and disease conditions, as well as laboratory, epidemiologic, and biostatistical method improvements as well as potential shifts in etiologies raise concerns about data comparability across large intervals of time.

It is quite likely that combinations of complementary markers might be required for different purposes. For instance, specific growth measures are used for population screening (e.g., height-for-age Z-scores to identify prevalence of stunting within given populations), and differ from those used for individual clinical management (i.e., growth monitoring). Another paradigm might be to use a highly sensitive screen for enteropathy, followed by more specific reflexive testing. An example in clinical medicine would be HIV screening followed by protein immunoblot (Western blot) assay for specific antigenic reactivity. The situation might, however, be more complicated for EED, as there is unlikely to be a unifactorial process or driver. Ideally, forthcoming research will identify a marker or markers that can serve related but differing functions. The marker or set of markers might also vary according to the function desired. For EED, we foresee multiple potential points of testing. First, in an individual, a biomarker could be used to detect early pathologic or functional changes, at an actionable stage before stunting and its consequences ensue. Second, a biomarker could be used to monitor clinical progression/recovery to determine intervention effectiveness. Third, a biomarker could be used to measure population prevalence to determine those populations where focused attention could mitigate morbidity and mortality from enteropathy and stunting.

It is possible that a marker of end-pathway processes common to non-EED forms of enteropathy will have more utility than EED-specific biomarkers. However, we wish to note that the “re-purposing” of biomarkers of other enteropathies will not result in an EED-specific marker.

This issue would be most relevant if there is geographic overlap between EED and other enteropathies of childhood, most notably celiac disease, as might be the case in South Asia as evidenced by recent studies from India [315].

Different approaches probably need to be tested or developed in parallel to find the optimal biomarkers for enteropathy, and in particular, to identify the optimal biomarkers to detect enteropathy that results in adverse consequences for the host (mostly likely stunting). Such strategies could consist of broad screening and biomarker discovery, using agnostic methodologies (e.g., proteomics, genomics, and metabolomics). It is important to note that biomarker discovery requires disciplined approaches, strong statistical support to identify attributable risk to a specific biomarker and to avoid false discoveries, as well as validation cohorts. Underlying any such discovery and validation is the obligation to be absolutely certain that the core data and specimens are of the highest quality. To date, cohorts have been small, specimens not uniformly collected and handled, and outcomes have been nebulous. These problems need to be remedied before applying novel analytics to the problem of enteropathy. However, the MAL-ED [316] and other cohorts might provide sufficient rigor and resources to begin to answer these questions.

If a perfect, or even an adequate single biomarker, cannot be found, an alternative might be the development of an “enteropathy index” which integrates a constellation of clinical symptoms and signs and/or set of markers. Variations of the “enteropathy index” might need to accommodate the processes of screening, diagnosing dysfunction in individual cases, and monitoring this condition and its consequences. For example, we divided the markers into eight categories, representing different enteric functions or expressions. An “enteropathy index” might be represented by “stacking” one marker from each of three or more categories such as lactoferrin (intestinal inflammation) + L:M (primarily permeability but mannitol excretion provides a measure of absorptive function) + lactose hydrogen breath test (digestion). We wish to note

that a recent publication by Kosek, et al. [317] provides data in support of this concept, proposing the use of three stool markers (neopterin, myeloperoxidase, and alpha-1-antitrypsin) reflective of inflammatory and permeability processes.

In considering data to be obtained from biopsy specimens, it will be important to balance information to be gleaned against the small risk related to the procedure. Based on current state of knowledge, it is not known when enteric functional deficits related to environmental enteropathy occur in temporal relation to histopathologic changes. When EED research does incorporate obtaining biopsies, such issues should be considered. Given the challenges in obtaining pediatric biopsies, any future EED biopsy research should include ascertainment of the relationship between histopathology and growth, in addition to utilization of histopathology as a “gold standard” by which less invasive biomarker performance is gauged.

Whichever of these biomarker discovery and validation strategies is employed, we urge rigorous adherence to standards for the design and reporting of biomarker studies. The “STAndards for the Reporting of Diagnostic accuracy studies” or “STARD” initiative provides an excellent framework and checklist for diagnostic accuracy studies, allowing readers to assess for potential bias in a study (internal validity) and to evaluate generalizability (external validity). The STARD initiative is akin to the CONSORT statement for randomized controlled trials [318] and provides an excellent structure for consideration for development of biomarker discovery study design and reporting guidelines [319].

In reviewing these studies, it appears that many tests were compromised by choice of analyte, and independent of processes, it is important, looking forward, to consider the body substance analyzed, and the benefits and drawbacks of each. These are summarized in Table 16.

Table 16. Analyte attributes.

Comments on attributes of analyte, relevant to current and future biomarker assessments.

Attribute	Blood	Urine	Stool	Breath	Biopsies	Saliva
Invasiveness ¹	High	Low	Low	Med	High	Med
Steady state assessment ²	Yes	Sometimes	Sometimes	Yes	Yes	Yes
Requires processing before freezing or long-term storage ³	Yes	No	No	No	Yes	No
Schedulable ⁴	Yes	No	No	Yes	Yes	Yes
Consent issues ⁵	Medium	Low	Low	Low	High	Low
Components of analyte that can be studied	Plasma, serum, white cells, red cells, host nucleic acids	Mostly urine	Stool, microbes	Gas	Bulk tissues, cells, microbes	Saliva, microbes
Technology to obtain ⁶	Low	Low	Low	Med	High	Med (if special swabs employed)
Biohazardous ⁷	High	Low	Low	Med	High	Low
Offensive ⁸	Med	Low	High	Med	Med	Med
Ability to collect and bring from home ⁹	No	Yes	Yes	No	No	Yes
Provenance ¹⁰	Good	Fair	Fair	Good	Good	Fair

¹ Invasiveness is defined as "low" if analyte is produced spontaneously by subject, "medium" if a special apparatus or collection device is needed to obtain the specimen and "high" if mucosa or integument is breached to obtain.

² Steady state is assessed ("yes") if the substance studied is likely to remain stable on an hour-to-hour basis, and not assessed ("no") if substance studied is capable of changing profoundly in a short interval.

³ Processing (e.g., physical separation by centrifugation of components) is generally performed before long-term storage.

⁴ Schedulable samplings are those that can be arranged and obtained in advance, i.e., are not generally produced on command by children.

⁵ Consent issues are "low" if the substance is spontaneously produced without any physical risk to the subject, "medium" if obtaining the substance has a degree of physical discomfort or anxiety risk such as a hematoma or pain from a phlebotomy, and "high" if there is any possibility of injury that could result in hospitalization.

⁶ "Low" technology describes minimally expensive materials that are easily obtainable, such as needles and syringes; "medium" cost pertains to special-order items, such as breath hydrogen bags; and "high" cost pertains to materials obtained in an inpatient setting using sophisticated technology, such as endoscopic and biopsy equipment.

⁷ Materials are considered to have "low" biohazard risk if the likelihood of contracting a life-threatening infection are nil if the substance is placed on an open cut or inhaled; "medium" if a serious but rarely life threatening infection would result if the substance is placed on an open cut or inhaled (breath is placed in this category because of the possibility of tuberculosis or other infections spread via respiratory droplets); and "high" if a life-threatening infection, most notably HIV, would result if the substance is placed on an open cut or inhaled.

⁸ Offensiveness relates to a general reaction to a worker's skin coming in contact with the substance accidentally.

⁹ Materials can be brought from home for analysis, with a modest amount of instruction and materials.

¹⁰ Provenance relates to low, medium, or high risk of accidental or intentional substitution by a research subject or patient or family.

As can be seen, each body substance has strengths and drawbacks in research studies and clinical practice, and much of the literature chose specimens to analyze based either on precedent or lack of suitable alternatives. In reviewing the tests, blood has many commendable attributes, but the requirement to breach the skin to obtain the specimen adds an extreme, and often limiting, constraint to sampling. We wish to propose for consideration the use of transcutaneous technology to evaluate blood components. These devices are quite successful in measuring blood gases (oxygen and carbon dioxide), bilirubin, and glucose, and their methodology might be adaptable to either endogenous targets (e.g., iron) or challenge substances (D-xylose), and therefore have a high likelihood of playing a useful role in detecting and monitoring enteropathy.

In all assessments, it is important to recognize that normal or reference values are highly abstract concepts. Often, normal values were obtained from studying cohorts of children in developed countries, and there are rarely opportunities to establish normal values based on putatively normal subjects in the same environment, and of the same genetic makeup, as the subjects studied for enteropathy. Unless and until methodology and sampling of subjects are used to establish such norms, it will be difficult to calibrate the magnitude of abnormality. This problem is compounded by definitional issues in pathology. While we believe that intestinal inflammation is quite likely undesirable, we cannot state with certainty that an inflamed mucosa is wholly deleterious for children in resource-poor settings. Specifically, it is possible that low-to-moderate-grade inflammation of the gut is a homeostatic mechanism in environments in which enteric pathogens are ubiquitous.

The obstacles we encountered in this review suggest opportunities to adhere to publication standards in future reports of biomarker research related to EED. We were largely hindered by inadequate definitions and reporting incompleteness. The STARD guidelines are a very good basis for subsequent reports, but might need to be adapted in consideration of the

biomarker field. Specifically, processes that are interrogated by the biomarker should be clearly described, if known, and if not known, then that should be stated. Rigorous statistics should be applied for two overriding reasons: first, it is critical that point estimates (including percent attribution where study design allows for calculation) and confidence intervals should be reported in addition to the less informative p-values; second, in the assessment of multiple biomarkers identified in agnostic discovery projects, it is important to incorporate statistically valid safeguards against false discovery via multiple comparisons.

We also believed that the authors of the texts we reviewed generally obtained more data than they reported, probably because of page limitations or focus of the paper. Very few of the publications reported data that were obtained exclusively in the context of biomarker development or evaluation; many of the reports produced biomarker data in the context of other studies. Also, there was sparse use of online repositories for storage of meta-data. As such, considerable opportunities to review relevant data for biomarkers might have been lost. Certainly, there are increasing capacities to archive ancillary data for interrogation by other investigators. We are encouraged by this trend and believe that complete data deposition should be considered a “best practice” of future enteropathy-biomarker projects. Furthermore, systematic deposition is best facilitated, for ease of input as well as extraction of data, by placement in the public domain.

In addition to more rigorous and transparent provision of study design and major findings, it is apparent that future research in the field will be handicapped by current journal publication constraints. Specifically, biomarker studies in human populations are likely to generate volumes of data that are well in excess of the word and table limits of standard journal articles. This is especially regrettable, because the cost of accruing human cohorts is increasing, and therefore there is an obligation to make the most of the data that are produced henceforth. We propose that a biomarker data repository be established, and used to advance

research in this field, and to be accessible for future data queries. We do acknowledge, however, that the establishment, preservation, and curation of data require resources that are often not provided to the projects in which the data are generated. Nonetheless, a standardized, disciplined, and logical deposition and maintenance of data while studies are performed will reduce the expense of archiving the material after the study. Such investments will enhance and expedite biomarker research and application for reducing the burden of stunting.