FDA Public Workshop on Non-microbial Biomarkers of Infection for *In Vitro* Diagnostic Use

The information and questions contained in this document are not binding and do not create new requirements or expectations for affected parties, nor is this document meant to convey FDA's recommended approaches or guidance. Rather, the information contained in this document offers background and the basis for discussions at the Public Workshop.

The use of non-microbial biomarkers as indicators of infection offers the promise of significantly improving antibiotic stewardship; however, clinical validation of these assays remains a significant challenge to test developers. The goal of this document is to facilitate discussion of specific clinical issues regarding the validation of non-microbial biomarkers for infection at the upcoming FDA Public Workshop on October 16, 2015. The Workshop will seek input from academia, government, industry, clinical laboratories, and other stakeholders regarding possible approaches to the questions raised below.

Although the workshop title refers to "Non-microbial *Biomarkers* of Infection," in the context of this document "biomarker" should be considered synonymous with "analyte." It is expected that future FDA documents will refer solely to 'analytes' when describing non-microbial biomarkers of infection to avoid any confusion with other potential uses or definitions of biomarkers, particularly those that might be adopted by the FDA Biomarker Qualification program.

"Non-microbial" is specifically used to differentiate between analytes that are not biologically tied to a specific microorganism or family of microorganisms, such as C-reactive protein (CRP) or Interleukin 6 (IL-6), and microbial associated analytes, such as PCR-based detection of *M. tuberculosis* or antigen-based detection of influenza. There may be some analytes where this distinction is blurred, but for the purpose of the upcoming Workshop we expect the differences between 'microbial' and non-microbial' markers to be clear and not a subject of discussion.

A. Background

The diagnosis of infection can be challenging in the absence of a documented microbial pathogen. Determining the etiology of even common infections may be challenging for myriad reasons, including (among others) an inability to obtain specimens from the site of infection (e.g., pneumonia or otitis media), prior or concurrent antibiotic use that may affect culture-based isolation, and the absence of widely available rapid diagnostics that can provide actionable information prior to empiric treatment. Even when potential pathogens are identified, the true etiology of infection may be uncertain due to colonization, e.g., patients with pharyngitis and group A streptococcus detected, or ventilated ICU patients with mixed sputum cultures.

Increased emphasis on antibiotic stewardship has heightened interest in non-microbial based biomarkers in settings where infection may be suspected but difficult or problematic to confirm.[1] Differentiating disease etiology by non-microbial analytes is not a new concept; analytes such as procalcitonin (PCT) and CRP have strong advocates and detractors for use in patients with suspected or documented infection. As a broad generalization, the use of non-microbial biomarkers for the diagnosis of infection falls into two categories:

1. Identifying etiology or patient condition, e.g., infection or non-infection:
    - Differentiating sepsis from SIRS, i.e., differentiating infection from non-infection as the etiology of cardinal signs of inflammation in critically ill patients.
    - Differentiating bacterial from viral or non-infectious causes of respiratory symptoms.
    - Distinguishing occult/early infection from non-infection, e.g., in premature neonates.
2. As a prognostic or predictive marker:
    - Discriminating the degree of illness or progression for septic patients, (e.g., sepsis from severe sepsis or septic shock).
    - Use as a prognostic marker for predicting outcome in severely ill patients (e.g., mortality).
    - Use for assessing the need for intervention, e.g., continuing or altering antibiotic therapy.

Progress in the development and validation of analytes for the diagnosis of infection has been hampered by uncertainty regarding clinically relevant performance standards as well as the diversity of approaches to clinical study design, comparator methods, and statistical analysis. Study design directly affects the interpretation of assay performance, and as such, is critical for validation of new diagnostic tests proposed to distinguish infection (and/or type of infection) from non-infection.

The summary discussed below focuses primarily on specific issues for two distinct intended uses for non-microbial biomarkers of infection: (1) use for distinguishing viral, bacterial, and non-infectious causes of respiratory symptoms; and (2) use for the diagnosis of sepsis. Both of these potential intended uses are the subject of much current research and ongoing device development. Non-microbial biomarker use for the prognosis of patients with sepsis will also be discussed, though this may not necessarily reflect only patients with bacterial infection.

[1] e.g., see Caliendo AM, Gilbert DN, Ginocchio CC et al. Better Tests, Better Care: Improved Diagnostics for Infectious Diseases, Clinical Infectious Diseases. 2013; 57 (suppl 3): S139-S170.

B. The Diagnosis of Viral versus Bacterial Infection for Respiratory Infections

The use of non-microbial biomarkers to discriminate between viral, bacterial, and non-infectious etiologies of symptoms in the setting of suspected respiratory infection has been a major focus of non-microbial biomarker research. There has been greater impetus for this research in the setting of rising antimicrobial resistance and increasing recognition of potential adverse effects from antimicrobials. Selected aspects of the development and validation of non-microbial biomarkers are briefly discussed below:

Intended Use

Non-microbial biomarkers submitted for FDA clearance or approval must have at least one specific intended use identified by the manufacturer. The Intended Use(s) for a specific assay can identify specific clinical syndromes, e.g., lower (LRTI) and/or upper respiratory infection (URI), can target specific patient populations (e.g., adults or children), or may be indicated for use in a clinical algorithm that includes microbiologically-based diagnostics.

Unless excluded from study by specific enrollment criteria, clinical studies in support of a new biomarker should enroll subjects that reflect the general demographics of the US population and these characteristics (e.g., gender, ethnicity, race) should be captured and analyzed as part of a manufacturer's submission. Age and health status are other potentially significant patient characteristics essential to understanding device performance and possible limitations with respect to clinical use.

Age: Although studies have generally excluded very young infants, age cutoffs and proposals for pooling data across age ranges have varied. It is important to demonstrate that results can be pooled across different age ranges (e.g., that performance in children can be considered similar to that of an elderly person), and if so, how this can be addressed in a statistically rigorous way given the pragmatic effects on sample size and study design.

'Healthy' versus 'immunocompromised' hosts: Studies of non-microbial biomarkers have generally excluded immunocompromised patients, both for the additional study design complexity this introduces and the difference in benefit/risk for immunocompromised hosts. In addition, the mechanism of action of immune-based analytes may lead to different performance estimates in distinct immunocompromised populations such as neutropenia or HIV. Patients without a classic immunocompromised status but preexisting risk factors such as COPD may also need to be considered as unique populations in certain situations.

Study Design

Assays may be designed to detect an increased likelihood of either bacterial or viral infection individually, or differentiate between the two (as well as non-infection) through measurement

of one or more analytes. Approaches that may be applicable for studying URI may not be possible for LRTI, e.g., withholding of antibiotic therapy, may be clinically acceptable as part of a randomized interventional clinical trial for URI but perhaps not for LRTI. Additional study design concerns may be introduced by specific markers, e.g., certain markers may only respond to specific infections and seek an intended use for the identified subset. In contrast to clinical trial sites where patients are enrolled and sampled, the selection of sites for laboratory testing will depend on how the device will be used; for point-of-care and Clinical Laboratory Improvement Amendments (CLIA)-waived tests the clinical sites should match the testing sites, but for moderately complex tests the testing sites may be separate from clinical sites. Patient self-testing for possible over the counter availability adds another interesting potential wrinkle to clinical study design.

Selected concerns regarding studies of non-microbial diagnostics for respiratory infections are briefly summarized below:

Comparator methods

Confirming an infectious etiology has proven difficult in clinical trials of upper and lower respiratory tract infection, due in part to the challenges of specifying an acceptable comparator for both entities.

a. Upper Respiratory Infection

Microbiological comparators, usually including PCR combined with more traditional antigen or culture based tests, are confounded by potential colonization (e.g., rhinovirus and group A streptococcus), particularly for pediatric populations. Inaccessibility of specimens for testing common URIs such as otitis media or sinusitis is an additional challenge to establishing a comparator. Due to the limitations of available comparator methods, approaches that integrate microbiological detection with other non-microbial biomarkers (e.g., WBC, CRP or PCT, even if not FDA cleared for this use) have been proposed but may themselves lack complete validation. Clinical diagnosis by blinded physician panels when diagnosis is still uncertain may also be problematic for URI.

b. Lower Respiratory Infection

Similar to URI, no universally recognized gold standard reference exists for determining the etiology of suspected LRTI. Comparator methods are limited due to the difficulty of routinely obtaining deep respiratory specimens; even with an unequivocal diagnosis of pneumonia based on imaging and physical examination, sputum specimens are often unavailable, particularly in the outpatient setting. The pattern of infiltrate on x-ray is not

sufficiently sensitive or specific to confirm a viral or bacterial etiology.[2] Composite approaches used in prior studies[3] have included respiratory cultures (i.e., sputum and BAL when available), blood culture (when available), viral/atypical bacteria PCR panels, and microbial biomarkers such as pneumococcal, legionella, and influenza antigens. However, regardless of the methods used, specific microbiological diagnoses may be lacking in a large majority of patients. Colonization (e.g., pneumococcus in outpatients or colonization with other organisms following hospitalization) may also be a concern, and there remains uncertainty as to whether viruses such as rhinovirus or bocavirus may be the cause of pneumonia. This has often led to the ad hoc incorporation of additional non-microbial biomarkers such as white blood cell count (WBC), CRP, and absolute neutrophil count to the construction of a comparator method for a new biomarker. In other instances overall clinical assessment/assignment by blinded physician review panels has been proposed.

Clinical Trial design

Device performance can be estimated using approaches such as: (1) prospective studies with a predefined comparator method; (2) prospectively archived samples in which samples and comparator information are systematically collected via a well-defined protocol and then subsequently tested; and (3) through retrospective studies. It is difficult to anticipate a situation where contrived specimens could be used as surrogates for actual clinical specimens since it is the host response that is being investigated. Generally, the scientific community has conducted most studies for non-microbial diagnostics using prospective trials, raising the significant challenge that even in large clinical trials, uncommon but clinically significant pathogens may be difficult to capture. This may be especially true for pathogens that may have geographical or temporal variation. Other concerns are pathogens that may be more prominent in certain age ranges or in patients with specific underlying conditions such as COPD. Trial designs can be static (i.e., evaluation at a given time point) or incorporate information from follow-up evaluations to yield additional clinical and diagnostic information. An interventional study may involve significant resources, but may be able to assess device accuracy in the setting of an uncertain comparator when the risk from no intervention (e.g., antibiotic therapy) is low.

Performance estimates

At present, in the absence of a gold standard 'clinical truth' reference method for either upper or lower respiratory infection, optimal composite comparator methods are uncertain.

[2] Ito I, Ishida T, Togashi K et al. Differentiation of bacterial and non-bacterial community-acquired pneumonia by thin-section computed tomography. European Journal of Radiology. 2009; 72(3):388-395.

[3] e.g., see Jain S, Self WH, Wunderink RG et al. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults. New England Journal of Medicine. 2015; 373:415-427.

This directly affects the evaluation of the performance estimates of a device for viral versus bacterial infection, since high overall accuracy may not be achievable in the setting of an imperfect comparator. As also noted earlier, another concern is whether results can be pooled across ages, pathogens, different underlying conditions (such as COPD), and host variables such as sex, race, or ethnicity.

Questions for Discussion Regarding the Diagnosis of Viral versus Bacterial Infection for Respiratory Infections

In view of the multiple issues regarding design, and evaluation of a non-microbial diagnostic that proposes to diagnose or differentiate bacterial, viral, and non-infectious causes of respiratory symptoms, we hope to have panel discussion on the following specific questions:

1. Please discuss appropriate comparator methods for studies of upper and lower respiratory infections – specifically address microbiological, clinical, radiological, and other components that should be included as part of a comparator method.
2. Please discuss what performance parameters (e.g., positive and negative agreements) should be considered appropriate for upper respiratory and lower tract infections.
3. Please comment if there should be minimum performance estimates for specific sentinel diseases that be confirmed with high sensitivity, e.g., influenza, group A streptococcus, etc.
4. Please comment whether the device Intended Use should exclude diagnoses where microbiological-based tests exist e.g., group A streptococcus or influenza. Please also address whether it would be clinically acceptable to exclude certain pathogens (e.g., rhinovirus) from the intended use of a device due to presumed colonization or (presumed) lack of eliciting an immune response?
5. Please address if results can be combined across all ages, or if performance should be confirmed independently in different age subgroups.
6. Please discuss if immunocompromised hosts or targeted subgroups (e.g., COPD, asthma), where consequences of inaccurate results may be greater, need to be studied in clinical trials or it is acceptable to include these as device limitations.

C. Biomarkers in the setting of sepsis

Many different non-microbial analytes have been proposed as potential biomarkers for the diagnosis of sepsis including cytokine/chemokines (e.g., IL-6), cell surface markers (e.g., CD64), receptor biomarkers and expression markers.[4] However, despite extensive study, few analytes

[4] e.g., see Faix JD. Biomarkers of sepsis. Critical Reviews in Clinical Laboratory Science. 2013; 50(1):23–36.

have gained acceptance in clinical practice. To date, procalcitonin is the only FDA-cleared analyte specifically targeting sepsis, although not specifically for the *diagnosis* of sepsis.

Sepsis is not a singular disease entity with a clearly defined pathophysiology relative to a non-septic state; as a clinically defined entity, sepsis encompasses heterogeneous etiologies and occurs in the setting of patients with very diverse underlying conditions. Also problematic is the distinction between septic states such as sepsis, severe sepsis, and septic shock. Although these terms are used as clinical descriptors with potential prognostic significance, the association between a biomarker and clinical progression has been problematic, even for biomarkers presumed to directly reflect the pathophysiology of sepsis, such as lactate. A significant percentage of septic patients may lack a microbiologically defined etiology, and the definition of sepsis itself may vary across studies.

Neonatal sepsis represents a unique entity that has been studied independently from sepsis in adults. The signs and symptoms of potential invasive bacterial infection are especially nonspecific in infants, who can rapidly progress from a mild prodrome to life-threatening illness. Prematurity is a major risk factor for invasive bacterial or viral infection, and the need for prolonged hospitalization and monitoring of this population has led to significant interest in markers that may predict infection in patients with nonspecific clinical symptoms or even when asymptomatic, with readily accessible physiological measures such as such heart rate variability studied as well as in vitro biomarkers. A significant concern with studying infants is that biomarker responses may differ substantially in the developing immune system, restricting extrapolation of the performance of non-microbial biomarkers from adults downward and mandating confirmation in this population.

Intended Use

One focus of non-microbial biomarkers for sepsis has been development of an analyte (or multivariate combination of analytes) that can distinguish between SIRS (Systemic inflammatory response syndrome) and sepsis, i.e., more rapidly identifying the likelihood of infection as the underlying etiology for a patient presenting as SIRS/sepsis. In this setting, the potential impact of non-microbial biomarkers on clinical practice is particularly significant. Physicians often balance the conflicting goals of antimicrobial stewardship with the importance of prompt antibiotic administration for critically ill patients, and highly accurate biomarkers could aid in this decision. In high-risk populations such as neonates or immunocompromised hosts, rapid recognition and empiric therapy when sepsis is suspected are essential to reduce patient morbidity and mortality.

In addition to the differentiation of infection and non-infection (analogous to the distinction between sepsis and SIRS), non-microbial biomarkers have also been studied for other potential

clinical applications in critically ill patients such as differentiating the degree of sepsis (following the 2001 American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference(ACCP/SCCM) definitions[5]), as prognostic markers of disease progression and/or mortality, and as measures of treatment efficacy (e.g., to guide duration of antibiotic therapy).

Study Design

Study design will vary depending on the specific intended use being sought; however, in all cases, the study design should confirm the proposed intended use by providing evidence of the safety and effectiveness of the device under its conditions of use. FDA has identified the following as general issues as of concern in the development of non-microbial diagnostics for infection:

Comparator methods:

This has been one of the more difficult aspects of study design for non-microbial markers of sepsis as microbiological confirmation is lacking for the majority of septic episodes or may be problematic in critically ill patients.

a. Differentiating Sepsis from SIRS

Consistent with the ACCP/SCCM definitions, this potential Intended Use serves as a surrogate for differentiating infectious from non-infectious causes of symptoms. The definition of clinical truth in this context shares challenges with analytes that seek to distinguish between viral versus bacterial infections. Although standard of care evaluation of sepsis often includes extensive microbiological evaluation, e.g., blood cultures, sputum and urine culture, *C. difficile* toxin, etc., for many, if not most subjects in clinical trials, an endpoint adjudication committee is necessary to integrate all clinical results (e.g., laboratory results, imaging, risk factors, etc.) to reach a final, if not definitive, assignment.

b. Prognosis/mortality

The serial assessment of non-microbial biomarkers during hospitalization has been studied as predictive of change in clinical status, e.g., 28 day mortality. To assess the independent contribution of the biomarker to predict outcome, these studies have also captured additional variables (e.g., SOFA score, age, WBC count, etc.) in regression analysis,[6] although prognosis may also be strongly affected by non-quantifiable variables

[5] Levy MM, Fink MP, Marshall JC et al: 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Critical Care Medicine. 2003; 31:1250–1256.

[6] e.g., see Mikkelsen, M, Miltiades AN, Gaieski DF et al. Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock. Critical Care Medicine. 2009; 37:1670-1677.

that can be broadly captured as representative of clinical status such as the need for prolonged ICU care.

c. Clinical intervention

Non-microbial biomarkers have also been proposed as an aid for guiding interventions, e.g., duration/discontinuation of antibiotic therapy. This may represent the most difficult Intended Use to study and is likely to require prospective controlled studies.[7]

Clinical Trial Design

Most clinical trials assessing performance for sepsis biomarkers have been conducted as prospective studies since unbiased archived specimen collections with complete clinical datasets are uncommon, although this may be more common going forward. The specific data collected and captured in case report forms is a critical aspect of these studies as the final study endpoint is often resolved by physician panels. Serial data collection is usually necessary since a diagnosis of sepsis versus SIRS may not be clear even by thorough review of all the information available at the initial workup, even though the non-microbial marker may be collected only at a single early time point.

Studies of biomarkers intended to guide treatment are particularly complex as careful enrollment criteria and definitions of response are essential, e.g., as when applied to specific disease entities such as pneumonia. In most respects design of these trials is more similar to drug trials than traditional device studies, and sophisticated means to increase study power (and reduce samples size) such as Response Adjusted for Duration of Antibiotic Risk methodology (RADAR)[8] have been proposed.

The patient population to be enrolled into a clinical trial could potentially vary with the specific biomarker, although to date clinical studies of sepsis biomarkers have broadly enrolled non-immunocompromised adult patients with SIRS or sepsis per the 2001 ACCP/SCCM definition. More narrow study inclusion criteria may affect the scope of the intended use in package labeling.

Performance estimates

Similar to non-microbial diagnosis of viral/bacterial respiratory infections, there is no gold standard comparator method for the diagnosis of sepsis versus SIRS; however, due to the

[7] e.g., see Christ-Crain M, Jaccard-Stolz D, Bingisser R et al. Effect of procalcitonin-guided treatment on antibiotic use and outcome in lower respiratory tract infections: cluster-randomised, single-blinded intervention trial. Lancet. 2004; 363:600-7.

[8] Evans SR, Rubin D, Follmann D et al. Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR). Clinical Infectious Disease. 2015; 61(5):800-6.

clinical implications of acting on a false negative result, high sensitivity may be necessary to yield actionable information from a single result. Alternatively, the value of a non-microbial diagnostic may not be greatest as an independent measurement at the onset of disease but for assessing over time the likelihood of sepsis or other clinical markers as the clinical course evolves. For prognostic variables, there has not been a clear precedent for assessing the level at which device results are clinically useful.

Questions for discussion regarding study of non-microbial biomarkers in the setting of sepsis

1. Please comment on the appropriate comparator method for the diagnosis of sepsis.
2. Please discuss general design principles for studies of non-microbial biomarkers to guide intervention, such as duration of antibiotic therapy.
3. Please discuss the information that should be captured in case report forms for use in the diagnosis of sepsis.
4. Please discuss what should be considered clinically acceptable performance characteristics (e.g., positive percent agreement, negative percent agreement) for sepsis biomarkers in the context of different potential Intended Uses.
5. Please address if results can be combined across all ages, or if performance should be confirmed independently in different age subgroups. Please also discuss if performance characteristics should be different for the diagnosis of infection in neonates and the best approach to establishing these standards.