Biomarkers of sepsis

John C. Marshall, MD; Konrad Reinhart, MD; for the International Sepsis Forum

Background: **A complex network of biological mediators underlies the clinical syndrome of sepsis. The nonspecific physiologic criteria of sepsis syndrome or the systemic inflammatory response syndrome do not adequately identify patients who might benefit from either conventional anti-infective therapies or from novel therapies that target specific mediators of sepsis. Validated biomarkers of sepsis may improve diagnosis and therapeutic decision making for these high-risk patients.**

Objectives: **To develop a methodologic framework for the identification and validation of biomarkers of sepsis.**

Methods: **A small group meeting of experts in clinical epidemiology, biomarker development, and sepsis clinical trials; selective narrative review of the biomarker literature.**

Results: **The utility of a biomarker is a function of the degree to which it adds value to the available clinical information in the domains of screening, diagnosis, risk stratification, and monitoring of the response to therapy. We identified needs for greater standardization of biomarker methodologies, greater methodologic rigor in biomarker studies, wider integration of biomarkers into clinical studies (in particular, early phase studies), and increased collaboration among investigators, pharmaceutical industry, biomarker industry, and regulatory agencies.**

Conclusions: **Biomarkers promise to transform sepsis from a physiologic syndrome to a group of distinct biochemical disorders. This transformation could aid therapeutic decision making, and hence improve the prognosis for patients with sepsis, but will require an unprecedented degree of systematic investigation and collaboration. (Crit Care Med 2009; 37:2290–2298)**

KEY WORDS: sepsis; biomarkers; infection; risk stratification; clinical trials

Sepsis is a complex syndrome resulting from the innate host response to invasive infection. Sepsis is considered to be severe when accompanied by clinically important derangements in physiologic organ system function. When this process jeopardizes tissue perfusion, septic shock is said to be present (1, 2). Severe sepsis and septic shock are common indications for admission to an intensive care unit and the leading cause of morbidity and mortality for critically ill patients (3, 4).

The widely used clinical definitions for sepsis belie the fact that the process is highly heterogeneous in its underlying causes and expression, and that the clinical management and the prognosis for recovery are equally variable. Patients with sepsis typically differ with respect to the inciting organism and focus on infection, and also vary with respect to optimal approaches to antibiotic selection and surgical source control. A definite microbiological diagnosis cannot be made in one third or more of patients with clinical manifestations of sepsis (5, 6). Furthermore, even when patients with similar bacteriologic or anatomical presentations of infection are considered, patterns of morbidity and the ultimate prognosis for recovery vary from one patient to the other and from one geographic region to the next (6).

Sepsis arises through the activation of an innate immune response, with changes in the expression and activity of thousands of endogenous mediators of inflammation, coagulation, and intermediary metabolism (7). Discrete components of this response, therefore, are attractive experimental targets of therapy (8). Yet, unresolved heterogeneity—in both the biology of illness and the clinical strategies used to support septic patients—has hampered the development and evaluation of therapies for sepsis. The syndromes of sepsis, severe sepsis, and septic shock are defined by nonspecific alterations in physiology, rather than by specific cellular processes that represent potential therapeutic targets. Sepsis, as currently defined, comprises a concept (that morbidity arises from the host response to infection) that is imperfectly translated into a clinical syndrome by the use of consensus-derived, and nonspecific clinical and laboratory variables (1, 2, 9). However, we currently lack the capacity to delineate distinct populations of patients with a discrete disease—a key prerequisite to enable the development of specific biologically rational therapies (10).

More than 100 distinct molecules have been proposed as useful biological markers of sepsis (11). It is not known which of these provides truly useful information nor even how such utility is best established. The convergence of a conviction that identifying useful biomarkers of sepsis would represent an important advance in sepsis research, with the recognition that no explicit approaches exist to accomplish this objective, prompted the International Sepsis Forum to convene the International Sepsis Forum Colloquium on Biomarkers of Sepsis in Toronto, Canada, October 28–30, 2005. Here, we reviewed emerging concepts in biomarker development and vali-

From the Li Ka Shing Knowledge Institute (JCM), the Keenan Research Centre, and the Departments of Surgery and Critical Care Medicine (JCM), St. Michael's Hospital, and the University of Toronto (JCM), Toronto, Ontario, Canada; and the Department of Anesthesiology and Intensive Care Medicine (KR), Friedrich-Schiller University, Jena, Germany.

Dr. Marshall has consulted for Spatial Diagnosis and has received honoraria from Eli Lilly. Dr. Reinhart has consulted for BRAHMS-Diagnostics and has received honoraria from BRAHMS-Diagnostics. Dr. Reinhart has stock options in BRAHMS-Diagnostics and in USIRS-Lab.

This report is based on an expert colloquium on biomarkers of sepsis convened by the International Sepsis Forum (ISF) in October 2005. The ISF is a not-for-profit organization dedicated to improving the management of critically ill patients with sepsis. Its activities are funded by unrestricted grants from a variety of sponsors, including Eli Lilly, GSK, Eisai, Takeda, Toray, Exponential Biotherapeutics, BRAHMS Diagnostica, Spectral Diagnostics, Biosite, and Biomerieux.

For information regarding this article, E-mail: marshallj@smh.toronto.on.ca

DOI: 10.1097/CCM.0b013e3181a02afc

Table 1. Uses of biomarkers

Screening
To identify patients at increased risk of adverse outcome to inform a prophylactic intervention, or further diagnostic test
Diagnosis
To establish a diagnosis to inform a treatment decision, and to do so more reliably, more rapidly, or more inexpensively than available methods
Risk stratification
To identify subgroups of patients within a particular diagnostic group who may experience greater benefit or harm with therapeutic intervention
Monitoring
To measure response to intervention to permit the titration of dose or duration of treatment
Surrogate end point
To provide a more sensitive measure of the consequences of treatment that can substitute for a direct measure of a patient-centered outcome

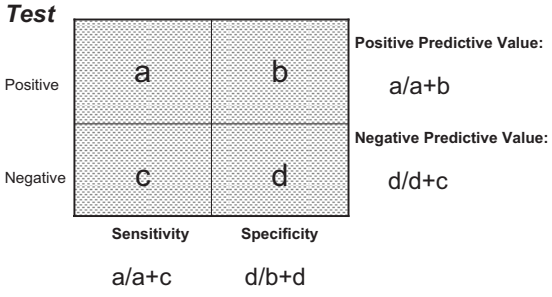

Figure 1. Determination of the sensitivity, specificity, positive predictive value, and negative predictive value of a diagnostic test. The positive likelihood ratio is calculated as the sensitivity/1-specificity, and the negative likelihood ratio as 1-sensitvity/specificity. Positive likelihood ratios over the range of values of the diagnostic test is represented by the receiver operating characteristics curve, the area under the curve being a reflection of the accuracy of the test across a range of values.

dation, and integrated these with recommendations of the colloquium for future research priorities.

A biomarker is "... a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (12), or more simply, as a "quantifiable measurement(s) of biological homeostasis that defines(s) what is 'normal,' therefore providing a frame of reference for predicting or detecting what is 'abnormal'" (13). The utility of a biomarker lies in its capacity to provide timely information beyond that which is readily available from routine physiologic data and clinical examination. This additional information may provide insight into the pathogenesis or prognosis of a disease process and also aid in a therapeutic decision; further, it may facilitate titrating therapy or monitoring the response to intervention (Table 1).

The performance of a marker can be evaluated by its sensitivity (its ability to detect a disease in patients in whom the disease is truly present) and its specificity (its ability to rule out a disease in patients in whom it is truly absent) (Fig. 1). Ideally, these properties should be established in a laboratory model of the disease of interest to ensure that the test performs reliably and accurately under optimally controlled and reproducible circumstances. However, this is rarely possible, and we typically use a test in an ambiguous clinical setting where the diagnosis is uncertain. Furthermore, our interest lies in knowing how likely it is that if a test result indicates that the disorder is absent, the disorder is truly absent, and if the test result indicates that the disorder is present, the disorder is truly present. This information is embodied in test properties called the *negative and positive predictive values*, respectively (Fig. 1).

The discriminatory value of a diagnostic test is reflected in the ratio of the measured positives in patients in whom the disease is truly present (sensitivity), to the measured positives in patients in whom the disease is truly absent (1-specificity). This ratio is known as the *likelihood ratio*. A likelihood ratio of 1 indicates that the test performs no better than random chance; the utility of the test increases as the likelihood ratio becomes greater or less than 1, reflecting positive or negative likelihood ratios, respectively. Likelihood ratios have a distinct advantage over metrics, such as sensitivity, specificity, positive predictive value, and negative predictive value, because they can be calculated for multiple levels of the test. Furthermore, likelihood ratios at differing values of the test can be plotted graphically to produce a receiver operating characteristics curve: superior performance is reflected in a higher value for the area under the receiver operating characteristics curve.

Uses of Biomarkers

A biomarker may serve one or more of five overlapping roles (Table 1). It may identify a patient with an increased probability of having a disease or an adverse outcome, and also serve as a screening test, detecting risk before the development of a clinical disease, or identifying a patient who might warrant evaluation by a more definitive, but expensive or invasive diagnostic test. For example, the presence of occult blood in the stool suggests the possibility of colorectal cancer and identifies a subpopulation of patients who are more likely to benefit from colonoscopy (14), whereas in a population of critically ill patients, an elevated white blood cell count might trigger a more intensive investigation to identify a focus on infection.

Second, a biomarker may identify the presence of a pathologic state or process and also establish a diagnosis. Added value arises from the capacity of the marker to establish (or rule out) the diagnosis more reliably, more rapidly, or more inexpensively than other available measures.

Third, a biomarker may aid in risk stratification—the parsing of an heterogeneous population of patients with a disease into a more homogeneous one, in whom the potential for benefit with a given therapeutic intervention is greater. For women with breast cancer, for example, estrogen receptor positivity defines a population of patients who might benefit from hormonal manipulation, whereas the presence of the human epithelial growth factor receptor 2 biomarker identifies a subpopulation of patients who might benefit from herceptin (15). Risk stratification implies differential prognosis, the identification of patients who are not only more likely to have either a favorable or an unfavorable outcome but

Table 2. Sources of variability in the measurement of cytokines in sepsis

Assay methodology
Bioassay
Immunoassay: enzyme-linked immunosorbent assay, bead, radioimmunoassay
Reference standards
Site of sampling
Cellular: membrane, cytoplasmic, nuclear
Regional (e.g., bronchoalveolar lavage fluid)
Systemic (blood): whole blood, plasma, leukocytes, leukocyte subsets
Requirement for activation
E.g., protein C, transforming growth factor β
Presence of soluble receptors, circulating inhibitors, and carrier proteins

also who are more or less likely to respond to a particular therapy.

Fourth, a biomarker whose levels change dynamically as the patient responds (or fails to respond) to treatment provides the clinician with information to monitor the response to intervention and also to titrate or modify the therapeutic intervention. In the intensive care setting, glucose levels provide information that allows the clinician to titrate insulin therapy, whereas the activated partial thromboplastin time guides the optimal dosing of heparin.

Finally, if changes in the level of the biomarker can be shown to correlate consistently with clinically important patient outcomes, then the marker may find a role as a surrogate outcome measure for preventive or therapeutic interventions, particularly during an early phase clinical research (16). Surrogate end points are outcomes that substitute for direct measures of how a patient feels, functions, or survives (17). Surrogate end points include physiologic or laboratory variables (for example, blood pressure as a surrogate end point for stroke) or measures of subclinical disease (e.g., degree of atherosclerosis on coronary angiography as a surrogate end point for the risk of myocardial infarction or cardiac death). The substitution of a biomarker as a surrogate end point for patient-important outcomes is attractive when the biomarker can be measured earlier, more easily, more frequently, with higher precision or with less confounding by competing risks or other therapies. To be valid for this purpose, the biomarker must not only correlate with the patient-important outcome but also must capture, to the greatest possible extent, the net effect of the intervention on the patient-important outcome (18). We will come back to this issue later.

Approaches to the Identification of Biomarkers

Potential biomarkers are identified by several approaches. First, a marker may be selected on the basis of a biologically compelling association with a disease state or a candidate therapeutic intervention. In patients with sepsis, for example, circulating endotoxin might be a marker of a patient with Gram-negative infection and would be anticipated to identify a patient who might benefit from treatment with an agent that neutralizes endotoxin. Conversely, circulating tumor necrosis factor (TNF), or a downstream molecule such as interleukin-6 whose release is induced by TNF, is an intuitively attractive marker of a patient who might benefit from a therapy that neutralizes TNF.

A biomarker may also be identified serendipitously on the basis of an apparent association with a disease in the absence of a biologically plausible link. Procalcitonin (PCT) levels, for example, were observed to be increased in patients with infection and to drop in response to adequate antibiotic therapy (19).

Finally, biomarkers may be identified by using unbiased approaches in which large numbers of molecular species are assayed using microarray or proteomic approaches to identify those species that are differentially expressed in the population of interest. The advent of techniques to simultaneously detect changes in many thousands of molecular species—messenger RNA transcripts using microarrays (20) or peptides in biological samples using mass spectrometry approaches (21)—has opened the possibility of genome-wide screening for candidate biomarkers. At the same time, the sheer volume of information generated creates enormous challenges in bioinformatics. The challenge is compounded by the fact that leukocyte transcriptomics and plasma proteomics in systemic inflammatory response syndrome yield discordant results if the circulating cell pool is not the source of the proteomic response (22). The selection of appropriate controls is crucial to ensure that changes detected are the consequence of the process of interest, and not, for example, of an unrelated clinical intervention such as transfusion.

Biomarker Validation

The validation of a biomarker requires consideration of three discrete domains of the biomarker's performance:

- Demonstration that the assay truly measures a particular molecular species, or its relevant biological activity
- Demonstration that measurement of the biomarker discriminates patients with a disease from those who are free from the disease, and/or stratifies patients having a disease on the basis of their risk of adverse outcome
- Demonstration that measurement of the biomarker can inform a clinical decision that leads to improved patient outcomes

What Does the Assay Measure?

Conclusions from cohort studies of biomarker levels are often discordant, reflecting the multiple sources of variability that may arise through subtle differences in assay methodology, reagents, and site of sampling, and the confounding effects of carrier proteins and circulating inhibitors (Table 2).

For example, endotoxin from Gram-negative bacteria is commonly present in the circulation of patients with both infectious and noninfectious acute illnesses (23, 24). The classic assay for endotoxin has been the Limulus amebocyte lysate assay, a bioassay based on the capacity of endotoxin to induce coagulation of a lysate from the hemolymph of the horseshoe crab, *Limulus polyphemus* (25). The reaction is not specific to endotoxin, but can also be activated by other microbial products, particularly components of fungal cell walls (26). Conversely, endogenous plasma proteins inhibit the reaction, reducing the reliability of the assay in protein-containing biological fluids (27). Furthermore, endotoxin *in vivo* is transported bound to a specific carrier protein, and its activity inhibited by other endogenous proteins, thus even when endotoxin is detected, it may not be biologically active. A bioassay based on the priming of neutrophil respiratory burst activity by complexes of endotoxin and antiendotoxin antibody has recently been reported (24); its utility and limitations still have to be established.

Comparable challenges arise in the assay of host-derived mediators such as TNF. Circulating TNF can be detected by assay of immunoreactive protein (by enzyme-linked immunosorbent assay or using multiplex bead array technology [28]) or assay of circulating bioactivity (the L929 or Walter and Eliza Hall Institute cell cytotoxicity assays). Quantitative data obtained from immunoassays typically

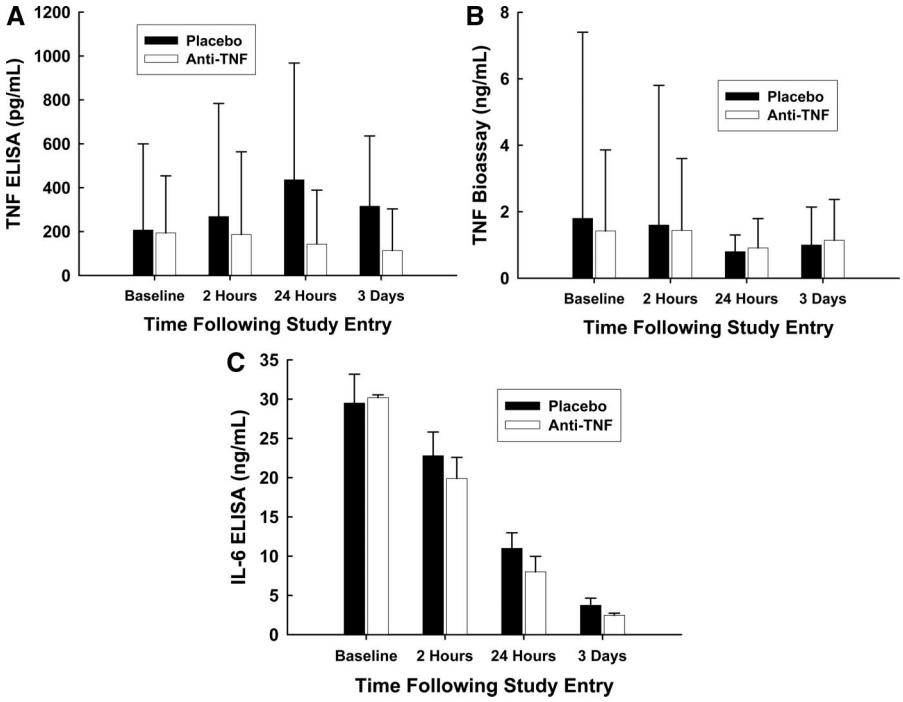

Figure 2. Cytokine data from a multicenter study of a monoclonal antibody to tumor necrosis factor (*TNF*) in patients with septic shock (31). Although the antibody reduced levels of immunoreactive TNF (*A*), it did not reduce the levels of bioactive TNF (*B*), and failed to impact levels of the downstream cytokine, interleukin (*IL*)-6 (*C*). Thus, the absence of a mortality benefit may reflect the use of a biologically inactive therapy. *ELISA*, enzyme-linked immunosorbent assay.

vary with the supplier of the assay and the nature of the antibody and reference standards used (29, 30). Furthermore, the association of immunoreactive protein with biological activity is inconsistent. In a phase III trial of a monoclonal antibody to TNF (31), treated patients were found to have increased circulating levels of the therapeutic antibody and reduced levels of immunoreactive TNF; however, the levels of TNF bioactivity were comparable in the two populations (Fig. 2). It is further uncertain whether the biologically active TNF is that found in the circulation, or that which remains expressed on the cell surface (32), how TNF bioactivity might be modified through the concomitant shedding of cell surface receptors (33), and whether levels in the circulation are relevant to disease processes that are effected in the microenvironment of specific tissues.

Biomarker Prevalence in an At-Risk Population

In patients with sepsis, evaluation of a biomarker hinges on appropriate consideration of two issues:

1. What is the disease, and what is the gold standard diagnostic criterion against which the marker will be tested?
2. What is the appropriate control group in which the disease is considered to be absent?

Both pose daunting challenges.

The evaluation of a novel diagnostic marker requires its comparison with an existing measure, colloquially known as the *gold standard*. This might be a test that definitively establishes the presence of a disease—malignant cells in a histologic specimen from a lymph node establishing the presence of metastases or viable microorganisms in a lung biopsy establishing the diagnosis of pneumonia. More often, however, the diagnostic gold standard is one that has established its authority by use over time—for example, S-T elevation on an electrocardiogram as the standard for acute myocardial infarction, or quantitative cultures of the urine as the standard for a urinary tract infection. However, there is no comparable standard for the diagnosis of sepsis. First, sepsis is a concept—that of disease arising from the host response to infection—rather than a measurable pathologic process. Second, that concept is a complex one that hinges on documentation of both infection and a response. Third, that response is nonspecific, defined by consensus criteria that emphasize physiologic changes in vital parameters that are common to a number of disparate processes.

In the absence of an objective pathologic gold standard, the definition of sepsis depends not on arbitrary clinical criteria but on the specific decision that is to be made. If the decision is to initiate antibiotic therapy, the definition, and therefore the diagnostic criteria, must reflect the presence or absence of infection, and the identity of the infecting microorganism. If the decision is to use activated protein C, the question of interest is the presence or absence of deficient protein C activity, or the potential to respond to supplementation. The gold standard against which a biomarker of sepsis is evaluated must be defined with reference to the clinical decision that the marker will inform, and will also vary with the biomarker.

The ideal control group is one that is similar in all readily measurable characteristics to the population that the biomarker defines, but whose outcome without intervention differs from that of the population in which the biomarker is present (34). In other words, the populations should reflect true diagnostic uncertainty. The apparent utility of a biomarker or diagnostic test can be overestimated if the control group is systematically different from the population in whom the marker is studied (for example, if the controls are healthy laboratory volunteers), a form of bias termed *spectrum bias* (35).

Impact of Biomarker-Directed Decision Making on Clinical Outcomes

The most compelling validation of the biomarker performance is demonstration that the biomarker differentially identifies patients who experience benefit from a particular intervention. For example, demonstration of microsatellite stability in colonic cancers identifies a subgroup of patients who will benefit from treatment with 5-fluorouracil (36), whereas expression of the human epithelial growth factor receptor 2 on breast cancer cells identifies a population of patients whose survival can be prolonged by the administration of herceptin (15).

Although documentation that a biomarker can effectively stratify patients who are candidates for therapy provides convincing

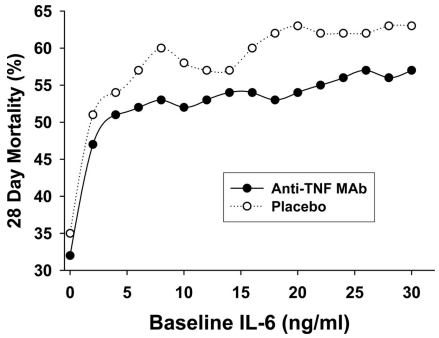

Figure 3. A large multicenter study randomized patients with severe sepsis to treatment with an antibody to tumor necrosis factor (*TNF*) or placebo, stratifying patients on the basis of baseline levels of interleukin (*IL*)-6 (38). Although the impact on mortality was statistically significant in an adjusted analysis of IL-6–positive patients, the incremental benefit over IL-6–negative patients was minimal. The selected cutoff of 1000 pg/mL may have been too low: analysis of the treatment effect over a range of IL-6 values revealed a greater separation of the mortality curves of placebo and anti-TNF–treated patients at higher levels of IL-6. *MAb*, monoclonal antibody.

evidence of clinical utility, it does require the performance of adequately powered studies to document benefit when the positive predictive value of the marker is being tested, or equivalence or noninferiority, when it is the negative predictive value that is the parameter of interest. Several trial designs are possible (37). An interventional study can be performed in which patients receiving one or other of two therapies are stratified on the basis of the marker of interest and differential efficacy in the two strata compared. This approach was used in a recently published study of an antibody to TNF in which patients were stratified by baseline interleukin-6 levels of more or less than 1000 pg/mL (38) (Fig. 3). Alternatively, patients may be randomized to have therapy directed by the marker, or provided without knowledge of the marker, and the differential consequences assessed. This latter design was used in a randomized trial assessing the utility of PCT levels in informing therapy in patients presenting to an emergency department with acute respiratory symptoms (39), and in a study of the use of PCT levels to direct the duration of antibiotic therapy for patients with community-acquired pneumonia (40).

Biomarkers as Surrogate Outcome Measures

A surrogate outcome measure is an "end point measured in lieu of some other so-called true end point" (41)—one that, while not itself a measure of a patient-centered outcome, reliably and consistently predicts a clinical outcome. Surrogate outcome measures are commonly used in clinical trials evaluating new therapies for infectious diseases, for example, clinical response rates in studies of antibiotic therapy for intra-abdominal infection (42), or viral load in studies of treatments for human immunodeficiency virus (43). Surrogate outcomes have generally been dismissed as appropriate outcomes for sepsis trials (44), for the critical care literature is replete with studies of interventions that have improved a physiologic or biochemical end point, but actually worsened survival (45–48). Conversely, wider use of surrogate measures could prove invaluable during an early phase clinical research in establishing proof of principle, in refining study entry criteria, and in establishing optimal dose and duration of therapy.

A valid surrogate measure must satisfy three criteria: it must predict disease progression, be affected by therapy, and respond to the same biological processes that are thought to mediate the clinically important outcome (18).

The optimal use of biomarkers as surrogates in informing the design of definitive clinical trials presupposes a valid and extensive understanding of the natural history of the biomarker in the population of interest, and how its levels are modified by therapeutic intervention. These data can then be integrated using meta-analytic techniques (49) to evaluate the capacity of a biomarker to predict a clinically important outcome. A methodology for evaluating the level of evidence that a given biomarker might serve as a reliable surrogate outcome measure has recently been proposed (50), but its utility in the assessment of biomarkers for diseases such as sepsis where mortality is considerable is unknown.

Biomarkers in Sepsis Research and Clinical Management

Although biomarkers of sepsis are not widely used in research or clinical practice, it is possible to evaluate the utility of approaches that are currently available. We will consider these within the framework of the nascent PIRO model—a model that seeks to stratify patients with sepsis in the four domains of predisposition, insult, response, and organ dysfunction (2).

Is the Patient at Increased Risk of Adverse Outcome?

Polymorphisms in innate immunity genes are common and result in significant interindividual variability in response to a given insult (51). Indeed, the risk of dying from an infectious disease is much more dependent on genetic than on environmental factors (52). Single nucleotide polymorphisms can be readily detected using polymerase chain reaction-based approaches (53), although such techniques are not currently optimized for rapid diagnosis. Single nucleotide polymorphism analyses are likely to be of greatest importance when an intervention targets a single key mediator. It has become apparent, for example, that the presence of a G->A polymorphism at −308 in the promoter region of the TNFα gene is associated with differential responsiveness to anti-TNF therapy in rheumatoid arthritis (54), although not in inflammatory bowel disease (55).

Current studies of genetic markers in sepsis extend beyond individual candidate gene variations and include genome-wide approaches, which promise insights into genes and their variations not commonly studied in sepsis. New techniques, such as genomic microarray assays, enable detection of hundreds of thousands of single nucleotide polymorphisms in a single patient, and also mostly reflect an individual's genomic uniqueness.

Is the Patient Infected?

Knowledge that a clinical syndrome of systemic inflammation is likely a consequence of invasive bacterial or fungal infection prompts the clinician to initiate appropriate empirical antibiotic therapy and to look for a locus of infection amenable to source control measures. Conversely, confidence that infection is unlikely to be present permits the clinician to withhold or discontinue antibiotics. The consequences are not only a reduction in costs but also minimization of the adverse consequences of therapy, including superinfection with organisms such as *Clostridium difficile* (56).

The commonly used laboratory parameter of leukocytosis has a very low sensitivity and specificity for the diagnosis of infection, with a likelihood ratio of 1.5; band counts have a similarly low diagnostic accuracy (57, 58). C-reactive protein levels provide greater diagnostic information than temperature elevation

Table 3. Summary recommendations

We recommend that steps be taken to standardize assays for the measurement of biomarkers of sepsis, and to identify and understand the sources of differences among techniques, for example, bioassays vs. immunoassays, cell-bound vs. free protein, target biomarkers vs. levels of circulating inhibitors, and the confounding effects of inhibitors and carrier proteins. Similarly, the performance of microarray analyses must be standardized with respect to platforms, composition of arrays, and methods of data analysis
We recommend that studies of biomarkers of sepsis be performed using rigorous methodologic approaches to characterize the added value provided by the maker. Such approaches can include inception cohort studies, case–control studies, or randomized clinical trials, analyzed using multivariable techniques to define the independent diagnostic or prognostic value of the marker
We recommend wider use of validated biomarkers to assist in the decision-making process in guiding the transition from early phase clinical research to definitive trials with clinically important end points
We urge that clinical trials be used as platforms to identify and validate potentially useful biomarkers of sepsis—both to evaluate drug efficacy and to generate knowledge on variability in populations and changes with the evolution of disease
Notwithstanding the insights that can be gained from intensive evaluation within clinical trials and pooling of data across studies, we perceive a clear need for one or more large, intensive, comprehensive international study to define the biochemical natural history of sepsis, and to determine the association of biomarkers with disease progression, prognosis, and response to treatment
We urge increased collaboration between companies involved in diagnostics, and those involved in therapeutics, as well as greater collaboration among industry, clinical investigators, and regulatory agencies to advance our understanding of biomarkers in sepsis

in the diagnosis of infection in critically ill patients (59). However, systematic reviews suggest that PCT is superior to C-reactive protein, as evidenced by greater sensitivity, a higher positive likelihood ratio, and a greater area under the receiver operating characteristics curve (60, 61). Bronchoalveolar lavage levels of soluble triggering receptor expressed on myeloid cells-1 have been reported to be particularly accurate in the diagnosis of pneumonia, with a likelihood ratio of 10.4 (62); however, this marker has not been evaluated as extensively as PCT or C-reactive protein and requires the performance of an invasive procedure. The absence of elevated levels of circulating endotoxin appears to rule out a diagnosis of invasive Gram-negative infection (24).

What is the Microorganism?

The identification of an infecting microorganism using conventional microbiological techniques is inherently slow. Gram's stain can provide general information on the presence of a microorganism; however, growth on culture media and subsequent identification by laboratory techniques are needed to define the species of the organism and exposure to antibiotic-impregnated discs to determine antibiotic sensitivity. As a consequence, a definitive microbiological diagnosis may not be available until several days after the onset of the septic episode.

Rapid assays based on spectroscopy or polymerase chain reaction are under development and hold the promise of being able to detect multiple bacterial species, and to provide precise information on the presence of particular strains (63, 64). They are available for clinical use in Europe; although they are sensitive to low levels of bacterial DNA, their correlation with clinically significant infection is uncertain, and they suffer from an inability to differentiate viable from nonviable organisms.

Will This Patient Benefit from Specific Adjunctive Therapy?

Rapid diagnosis of the presence of elevated levels of a specific target of an adjuvant treatment—TNF or endotoxin, for example—or of reduced levels of a critical factor for replacement therapy—protein C, antithrombin, or interferon gamma, for example—is a prerequisite for the rational use of expensive and potentially toxic therapies (65).

Initial hopes that circulating interleukin-6 might identify patients who would benefit from treatments directed against TNF have proven disappointing (38), although it is possible that a higher cutoff for the diagnostic test would provide greater diagnostic accuracy (Fig. 3). The hypothesis that a state of relative adrenal insufficiency—identified on the basis of response to an adrenocorticotropic hormone stimulation test—defines a high-risk population who will benefit from treatment with exogenous corticosteroids (66) has been challenged by a recently published European study (67). Although there is evidence that patients with biochemical evidence of disseminated intravascular coagulation experience greater therapeutic benefit from activated protein C than patients who do not have disseminated intravascular coagulation (68), a reliable and validated biomarker to guide the use of this agent is not currently available. Reduced expression of the major histocompatibility marker, human leukocyte antigen, D-related, has been proposed as a biomarker for patients who might benefit from treatment with recombinant interferon gamma (69); however, further study is needed.

Given the enormous complexity and redundancy of the innate immune response, it is entirely plausible that the optimal use of biomarkers will require their integration into panels involving a number of analytes (70). Yet, even this approach poses substantial challenges. Wang et al (71) found that a panel of ten biomarkers of cardiac risk provided significant prognostic information for the population, but when applied to individual patients to evaluate capacity to predict risk beyond that available by conventional methods, the incremental benefit was small. A strong association between one or more biomarkers and population outcome will only translate into a useful diagnostic test if the distributions of the marker in affected and unaffected patients overlap only minimally (72).

Conclusions and Recommendations

Although there is widespread enthusiasm for a future role for the widespread use of biomarkers to inform the optimal management of the patient with sepsis, the field at present is underdeveloped. This underdevelopment provides the basis for the recommendations of this colloquium, summarized in Table 3.

First, there is substantial variability in the performance characteristics of available assays, in the types of assays used, and in the reference standards used to validate the assays. This has led to divergent study findings and to often discordant conclusions regarding the implications of changes in the expression of the biomarker of interest. We identified a need for a greater standardization of as-

say methodologies and for systematic comparisons of differing assay systems to more precisely understand their differences. We further recommend that published reports of biomarkers provide detailed information about the assay platform used, its performance characteristics, and the methods used for calibration.

We urge that investigators engaged in the study of novel biomarkers of sepsis use methodologically rigorous research designs and avail themselves of the rapidly increasing body of literature on the optimal conduct of studies of novel biomarkers. Cohort designs should report more than a simple association between a marker and an adverse outcome, such as death, and should seek to define the additional information that measurement of the biomarker provides. For diagnostic markers, an explicit, rigorous, and blinded process of adjudication of the clinical state that the marker is believed to diagnose (for example, ventilator-associated pneumonia or progression to septic shock) should be used and reported. We also recommend interventional study designs can be put to more use to assess, either indirectly or directly, the impact of the marker on clinical outcome (37). Finally, we recommend that authors of studies of biomarkers for sepsis adhere to emerging guidelines for the reporting of diagnostic studies, as articulated by the Standards of Reporting of Diagnostic Accuracy initiative (73).

We perceive an important missed opportunity in the use of biomarkers to inform the development of an early stage clinical research of therapies for sepsis and encourage the wider use of biomarkers as a mechanism for *post hoc* stratification to detect differential therapeutic efficacy in discrete, and biologically plausible subgroups of patients, and as surrogate outcome measures to detect proof of concept and characterize the biochemical consequences of intervention. Surrogate outcome measures can be of critical importance in an early phase clinical research in defining optimal patient populations to receive an intervention and in titrating the optimal dose and duration of treatment.

We urge those involved in large-scale randomized trials of treatments for sepsis to incorporate the measurement of a panel of biomarkers into the trial design, both to aid in future decisions regarding the use of the agent in the clinical arena, and to enhance our understanding of the natural history of sepsis, and the effects of specific interventions on its biological evolution.

Although valuable insights into the diagnostic and prognostic role of specific biomarkers can be gained from small cohort studies and interventional studies, the future evolution of critical care practice would benefit greatly from an enhanced understanding of natural history, and the development of disease descriptions based on distinct patterns of biochemical derangement, rather than on the nonspecific physiologic consequences of these events. We see the need for a comprehensive biological natural history study on the course of critical illness, designed to characterize the biochemical evolution, to facilitate therapeutic stratification and staging, and to understand the interaction of the changes of acute illness with therapeutic interventions. Such a study would proceed through the analysis of a large, heterogeneous cohort of patients, recruited on the criteria of being acutely ill (rather than the more restrictive criteria of systemic inflammatory response syndrome or acute respiratory distress syndrome, for example) and could evaluate the extent to which biological patterns of illness correlate with clinical manifestations and further facilitate the development of a robust staging system such as that proposed in the PIRO model.

Finally, it is clear that progress in sepsis research will require much greater collaboration among international groups of investigators, and between academia, industry, and regulatory agencies. We see the need for investigator-driven initiatives to define a research agenda, and to coordinate efforts to identify, validate, and implement clinically useful biomarkers for the management of septic patients. We urge greater collaboration between therapeutics and diagnostic companies in evaluating the diagnostic roles of specific biomarkers, and evaluating their response to therapeutic intervention. And we see a need for ongoing interactions among clinicians, investigators, industry, and regulatory agencies to improve the diagnosis and management of a vulnerable population of patients.

The transformation of a clinical syndrome into one or more diseases that can be characterized biologically is a prerequisite to the development of effective therapies. We look forward to the continuing evolution of critical care practice from its current role of nonspecific organ support to a discipline whose focus is the treatment of distinct diseases whose pathophysiology we only dimly understand today, and whose diagnosis remains, at present, elusive.

ACKNOWLEDGMENTS

We thank Elaine Rinicker and Peter Mainprice for logistic support for the symposium. We also acknowledge the contribution of all who participated in the International Sepsis Forum Symposium on Biomarkers of Sepsis: Co-chairs: John C. Marshall, Konrad Reinhart; Participants: Edward Abraham, Djillali Annane, Thierry Calandra, Deborah J. Cook, R. Phillip Dellinger, Jean-Francois Dhainaut, Paul Hébert, Daren Heyland, Lyle Moldawer, Rui Moreno, Steven Opal, Frank Stüber, Jean-Louis Vincent, and Hans-Dieter Volk.

REFERENCES

1. Bone RC, Balk RA, Cerra FB, et al: Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference. American College of Chest Physicians/Society of Critical Care Medicine. *Chest* 1992; 101:1644–1655
2. Levy MM, Fink M, Marshall JC, et al: 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. *Crit Care Med* 2003; 34:1250–1256
3. Martin GS, Mannino DM, Eaton S, et al: The epidemiology of sepsis in the United States from 1979 through 2000. *N Engl J Med* 2003; 348:1546–1554
4. Angus DC, Linde-Zwirble WT, Lidicker J, et al: Epidemiology of severe sepsis in the United States: Analysis of incidence, outcome, and associated costs of care. *Crit Care Med* 2001; 29:1303–1310
5. Sands KE, Bates DW, Lanken PN, et al: Epidemiology of sepsis syndrome in 8 academic medical centers. *JAMA* 1997; 278:234–240
6. Vincent J-L, Sakr Y, Sprung CL, et al: Sepsis in European intensive care units: Results of the SOAP study. *Crit Care Med* 2006; 34: 344–353
7. Calvano SE, Xiao W, Richards DR, et al: A network-based analysis of systemic inflammation in humans. *Nature* 2005; 437: 1032–1037
8. Marshall JC: Such stuff as dreams are made on: Mediator-targeted therapy in sepsis. *Nat Rev Drug Discov* 2003; 2:391–405
9. Bone RC, Fisher CJ, Clemmer TP, et al: Sepsis syndrome: A valid clinical entity. *Crit Care Med* 1989; 17:389–393
10. Marshall JC: Rethinking sepsis: From concepts to syndromes to diseases. *Sepsis* 1999; 3:5–10
11. Marshall JC, Vincent J-L, Fink MP, et al:

Measures, markers, and mediators: Towards a staging system for clinical sepsis. *Crit Care Med* 2003; 31:1560–1567
12. Biomarkers Definitions Working Group: Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001; 69:89–95
13. Dalton WS, Friend SH: Cancer biomarkers—An invitation to the table. *Science* 2006; 312:1165–1168
14. Fisher JA, Fikry C, Troxel AB: Cutting cost and increasing access to colorectal cancer screening: Another approach to following the guidelines. *Cancer Epidemiol Biomarkers Prev* 2006; 15:108–113
15. Duffy MJ: Predictive markers in breast and other cancers: A review. *Clin Chem* 2005; 51:494–503
16. Weir CJ, Walley RJ: Statistical evaluation of biomarkers as surrogate endpoints: A literature review. *Stat Med* 2006; 25:183–203
17. Bucher H, Guyatt G, Cook D, et al: Surrogate outcomes. *In:* The Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Guyatt G, Rennie D, (Eds). Chicago, IL, AMA Publications, 2002
18. Prentice RL: Surrogate endpoints in clinical trials: Definition and operational criteria. *Stat Med* 1989; 8:431–440
19. Assicot M, Gendrel D, Carsin H, et al: High serum procalcitonin concentrations in patients with sepsis and infection. *Lancet* 1993; 341:515–518
20. Feezor RJ, Cheng A, Paddock HN, et al: Functional genomics and gene expression profiling in sepsis: Beyond class prediction. *Clin Infect Dis* 2005; 41(Suppl 7):S427–S435
21. Liu T, Qian WJ, Gritsenko MA, et al: High dynamic range characterization of the trauma patient plasma proteome. *Mol Cell Proteomics* 2006; 5:2167–2174
22. Tomic V, Russwurm S, Moller E, et al: Transcriptomic and proteomic patterns of systemic inflammation in on-pump and off-pump coronary artery bypass grafting. *Circulation* 2005; 112:2912–2920
23. Opal SM, Scannon PJ, Vincent JL, et al: Relationship between plasma levels of lipopolysaccharide (LPS) and LPS-binding protein in patients with severe sepsis and septic shock. *J Infect Dis* 1999; 180:1584–1589
24. Marshall JC, Foster D, Vincent J-L, et al: Diagnostic and prognostic implications of endotoxemia in critical illness: Results of the MEDIC study. *J Infect Dis* 2004; 190:527–534
25. Levin J, Bang FB: Clottable protein in *Limulus*: Its localization and kinetics of its coagulation by endotoxin. *Thromb Diath Haemorrh* 1968; 19:186–197
26. Mitazaki T, Kohno S, Mitsutake K, et al: (1–>3)-beta-D-glucan in culture fluid of fungi activates factor G, a limulus coagulation factor. *J Clin Lab Anal* 1995; 9:334–339
27. Roth RI, Levin FC, Levin J: Optimization of detection of bacterial endotoxin in plasma with the limulus test. *J Lab Clin Med* 1990; 116:153–161
28. Elshal MF, McCoy JP: Multiplex bead array assays: Performance evaluation and comparison of sensitivity to ELISA. *Methods* 2006; 38:317–323
29. Sweep FC, Fritsche HA, Gion M, et al; EORTC-NCI Working Group: Considerations on development, validation, application, and quality control of immuno(metric) biomarker assays in clinical cancer research: An EORTC-NCI working group report. *Int J Oncol* 2003; 23:1715–1726
30. Lash GE, Scaife PJ, Innes BA, et al: Comparison of three multiplex cytokine analysis systems: Luminex, SearchLight and FAST Quant. *J Immunol Methods* 2006; 309: 205–208
31. Abraham E, Anzueto A, Gutierrez G, et al: Double-blind randomised controlled trial of monoclonal antibody to human tumour necrosis factor in treatment of septic shock. *Lancet* 1998; 351:929–933
32. Pellegrini JD, Puyana JC, Lapchak PH, et al: A membrane TNFα/TNFR ratio correlates to MODS score and mortality. *Shock* 1996; 6:389–396
33. McDermott MF: TNF and TNFR biology in health and disease. *Cell Mol Biol* 2001; 47: 619–635
34. Montori VM, Guyatt GH: Summarizing studies of diagnostic performance. *Clin Chem* 2003; 49:1783–1784
35. Lijmer JG, Mol BW, Heisterkamp S, et al: Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; 282: 1061–1066
36. Ribic CM, Sargent DJ, Moore MJ, et al: Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med* 2003; 349:247–257
37. Sargent DJ, Conley BA, Allegra C, et al: Clinical trial designs for predictive marker validation in cancer treatment trials. *J Clin Oncol* 2005; 23:2020–2027
38. Panacek EA, Marshall JC, Albertson TE, et al: Efficacy and safety of the monoclonal anti-TNF antibody $F(ab')_2$ fragment afelimomab in patients with severe sepsis stratified by IL-6 level. *Crit Care Med* 2004; 32:2173–2182
39. 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; 363:600–607
40. Christ-Crain M, Stolz D, Bingisser R, et al: Procalcitonin guidance of antibiotic therapy in community-acquired pneumonia—A randomized trial. *Am J Respir Crit Care Med* 2006; 174:84–93
41. Wittes J, Lakatos E, Probstfield J: Surrogate endpoints in clinical trials: Cardiovascular diseases. *Stat Med* 1989; 8:415–425
42. Solomkin JS, Reinhart HH, Dellinger EP, et al: Results of a randomized trial comparing sequential intravenous oral treatment with ciprofloxacin plus metronidazole to imipenem cilastatin for intra-abdominal infections. *Ann Surg* 1996; 223:303–315
43. Walmsley S, Bernstein B, King M, et al: Lopinavir-ritonavir versus nelfinavir for the initial treatment of HIV infection. *N Engl J Med* 2002; 346:2039–2046
44. Marshall JC, Vincent JL, Guyatt G, et al: Outcome measures for clinical research in sepsis: A report of the 2nd Cambridge Colloquium of the International Sepsis Forum. *Crit Care Med* 2005; 33:1708–1716
45. Hebert PC, Wells G, Blajchman MA, et al: A multicentre randomized controlled clinical trial of transfusion requirements in critical care. *N Engl J Med* 1999; 340:409–417
46. Takala J, Ruokonen E, Webster NR, et al: Increased mortality associated with growth hormone treatment in critically ill adults. *N Engl J Med* 1999; 341:785–792
47. Brower RG, Matthay MA, Morris A, et al: Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med* 2000; 342: 1301–1308
48. Lopez A, Lorente JA, Steingrub J, et al: Multiple-center, randomized, placebo-controlled, double-blind study of the nitric oxide synthase inhibitor 546C88: Effect on survival in patients with septic shock. *Crit Care Med* 2004; 32:21–30
49. Daniels MJ, Hughes MD: Meta-analysis for the evaluation of potential surrogate markers. *Stat Med* 1997; 16:1965–1982
50. Lassere MN, Johnson KR, Boers M, et al: Definitions and validation criteria for biomarkers and surrogate endpoints: Development and testing of a quantitative hierarchical levels of evidence schema. *J Rheumatol* 2007; 34:607–615
51. Holmes CL, Russell JA, Walley KR: Genetic polymorphisms in sepsis and septic shock: Role in prognosis and potential for therapy. *Chest* 2003; 124:1103–1115
52. Sorenson TI, Nielsen GG, Andersen PK, et al: Genetic and environmental influences on premature death in adult adoptees. *N Engl J Med* 1988; 318:727–732
53. Chen X, Sullivan PF: Single nucleotide polymorphism genotyping: Biochemistry, protocol, cost and throughput. *Pharmacogenomics J* 2003; 3:77–96
54. Seitz M, Wirthmuller U, Moller B, et al: The -308 tumour necrosis factor-{alpha} gene polymorphism predicts therapeutic response to TNF{alpha}-blockers in rheumatoid arthritis and spondyloarthritis patients. *Rheumatology (Oxford)* 2007; 46:93–96
55. Zipperlen K, Peddle L, Melay B, et al: Association of TNF-alpha polymorphisms in Crohn disease. *Hum Immunol* 2005; 66: 56–59
56. McDonald LC, Killgore GE, Thompson A, et al: An epidemic, toxin gene-variant strain of *Clostridium difficile*. *N Engl J Med* 2005; 353:2433–2441
57. Davis BH, Bigelow NC: Comparison of neutrophil CD64 expression, manual myeloid immaturity counts, and automated hematol-

ogy analyzer flags as indicators of infection or sepsis. *Lab Hematol* 2005; 11:137–147
58. Bogar L, Molnar Z, Kenyeres P, et al: Sedimentation characteristics of leucocytes can predict bacteraemia in critical care patients. *J Clin Pathol* 2006; 59:523–525
59. Povoa P, Coelho L, Almeida E, et al: C-reactive protein as a marker of infection in critically ill patients. *Clin Microbiol Infect* 2005; 11:101–108
60. Simon L, Gauvin F, Amre DK, et al: Serum procalcitonin and C-reactive protein levels as markers of bacterial infection: A systematic review and meta-analysis. *Clin Infect Dis* 2004; 39:206–217
61. Uzzan B, Cohen R, Nicolas P, et al: Procalcitonin as a diagnostic test for sepsis in critically ill adults and after surgery or trauma: A systematic review and meta-analysis. *Crit Care Med* 2006; 34:1996–2003
62. Gibot S, Cravoisy A, Levy B, et al: Soluble triggering receptor expressed on myeloid cells and the diagnosis of pneumonia. *N Engl J Med* 2004; 350:451–458
63. Peters RP, Van Agtmael MA, Danner SA, et al: New developments in the diagnosis of bloodstream infections. *Lancet Infect Dis* 2004; 4:751–760
64. Loy A, Bodrossy L: Highly parallel microbial diagnostics using oligonucleotide microarrays. *Clin Chim Acta* 2006; 363:106–119
65. Cohen J, Guyatt G, Bernard GR, et al: New strategies for clinical trials in patients with sepsis and septic shock. *Crit Care Med* 2001; 29:880–886
66. Annane D, Sebille V, Charpentier C, et al: Effect of treatment with low doses of hydrocortisone and fludrocortisone on mortality in patients with septic shock. *JAMA* 2002; 288: 862–871
67. Sprung CL, Annane D, Keh D, et al: Hydrocortisone therapy for patients with septic shock. *N Engl J Med* 2008; 358:111–124
68. Dhainaut JF, Yan SB, Joyce DE, et al: Treatment effects of drotrecogin alfa (activated) in patients with severe sepsis with or without overt disseminated intravascular coagulation. *J Thromb Haemost* 2004; 2:1924–1933
69. Docke WD, Randow F, Syrbe U, et al: Monocyte deactivation in septic patients: Restoration by IFN-gamma treatment. *Nat Med* 1997; 3:678–681
70. Ulloa L, Tracey KJ: The "cytokine profile": A code for sepsis. *Trends Mol Med* 2005; 11: 56–63
71. Wang TJ, Gona P, Larson MG, et al: Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med* 2006; 355:2631–2639
72. Ware JH: The limitations of risk factors as prognostic tools. *N Engl J Med* 2006; 355: 2615–2617
73. Bossuyt PM, Reitsma JB, Bruns DE, et al: Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003; 138:40–44