“Rule out sepsis” remains a persistently frustrating diagnosis in neonatology. It a common diagnosis: in the US, where 400 000 to 600 000 newborns are evaluated with a complete blood count (CBC) and blood culture each year, it is probably the second most common neonatal diagnosis after “well baby” (1–3). No one disputes that babies with clear-cut clinical signs of infection should receive systemic antibiotic therapy. However, beyond these two facts, and despite individual nursery protocols somehow keeping the problem—more or less—under control, multiple disagreements exist. Discord is most pronounced in two areas. The first is what test or combinations of tests permit one to diagnose sepsis early enough and reliably enough. After all, by the time a newborn shows clear-cut signs of infection, it is usually pretty late in the game (1).

The second concerns newborns with maternal risk factors who are initially asymptomatic: what constitute appropriate criteria for initiating antibiotic therapy in such infants (4–6)?

Faced with these questions, many investigators have attempted to find the diagnostic equivalent of a “magic bullet”, and a voluminous literature now exists advocating individual diagnostic tests or combinations of diagnostic tests, as well as combinations of a diagnostic test or tests with risk factors and/or clinical signs. All of these studies have yielded conflicting results, and the fact that these studies keep appearing in our journals highlights the fact that satisfactory solutions remain to be found.

In this issue, Chiesa et al. (7) report on a study that does not claim to have found the diagnostic equivalent of a magic bullet but that nonetheless breaks new ground and, because it is explicitly methodologic, has the potential to steer research on the early diagnosis of neonatal infections in a different direction. Unlike previous studies, whose emphasis has been to advocate this or that test or combination of tests, Chiesa et al. (7) took one step back and asked a very important question: what is the influence of illness severity on C-reactive protein (CRP), interleukin-6 (IL-6), and procalcitonin (PCT)? Importantly, they used two objective, validated measures of neonatal illness severity, the Score for Neonatal Acute Physiology (SNAP) and its perinatal extension (SNAP-PE) (8, 9). They found that CRP, IL-6, and PCT increase in the presence of bacterial infection and that their increases are independent of illness severity. They also found that illness severity has the potential to confound IL-6 concentrations in that, among newborns without bacterial infection, the higher the illness severity, the higher the IL-6 concentration after birth. Chiesa et al. (7) also confirmed that the sensitivity and specificity of these markers vary over time. Therefore, their use would require the use of specific cutoff values for each time point of evaluation over the first 48 h of life.

Does this mean that one should begin developing new nursery protocols to incorporate CRP, IL-6, and PCT? No, or not yet anyway, because Chiesa et al. (7), like others, found that the sensitivity of these markers is still acceptably low, in the range of 70–80% at birth, by far the most critical decision point when evaluating a newborn to rule out sepsis. It is true that the sensitivity of CRP increases over time, making serial measurements useful for those situations where one needs to decide how long to treat. However, by that point, most newborns will be asymptomatic and will have confirmed negative culture results. Thus, it is unlikely that busy clinicians (or the parents of an asymptomatic newborn with a negative culture at 24 h) would endorse the routine use of serial blood tests.

The real value of the work of Chiesa et al. (7) is that their approach suggests that it is time to move away from only seeking diagnostic magic bullets. Instead, we should also move toward better use—and better synthesis—of available information. Such an approach should include better use of the information provided by a given test as well as better use of information available before the performance of a given test. With respect to the information provided by a given test, it is essential to remember that the use of dichotomous cutoffs wastes information (10). In addition, with respect to other information, it is essential to remember that the posterior probability—and clinical value—of a given test result is highly contingent on the clinical situation and the prior probability (11).

These notions can be made concrete for any clinician by considering three clinical scenarios and reflecting on what the value of an increased CRP, IL-6, or PCT might be in each of them. The first is that of a newborn who presents with signs of a critical illness, e.g., respiratory failure, shock, and persistent seizures, within 1 or 2 h after birth. No one needs to measure biochemical markers or assign severity of illness scores to such infants. The only issue with respect to systemic antibiotics is how to get them in fast enough. The second scenario is that of a newborn who presents with equivocal signs, e.g., intermittent respiratory distress, desaturation, temperature instability, and blood pressure fluctuations. It is clear that, in these infants, the value of biochemical markers that are independent of illness severity could be extremely high. The third scenario is that of a newborn who is asymptomatic but who has risk factors (e.g., prematurity, premature rupture of membranes, prolonged rupture of membranes, maternal fever, or maternal carriage of group B streptococcus). The value of biochemical markers could also be high in these infants, but one would expect it to be strongly affected by the newborn’s a priori risk. One could also take this reflection further and consider what
the value of these tests would be as time passed and treatment or lack of treatment occurred.

At this point one can begin to adumbrate what the characteristics of a future cooperative study should be. First, as Chiesa et al. (7) have done, it is important to control for illness severity using objective criteria, a task made easier by the fact that the SNAP has been simplified (12). Second, rather than focusing on only the sensitivity and specificity of a given cutoff, the study should aim to define likelihood ratios for bacterial infection. That way clinicians could use the initial examination (which would categorize infants as asymptomatic, somewhat symptomatic, or critically ill) to define the prior probability for infection. Finally, the study should be structured so that a reasonable clinician could assess—given a baby’s risk factors and given the clinical examination—what the added value of CRP, IL-6, or PCT actually is. Only then can clinicians begin to make a rational choice as to when these tests should become a routine part of nursery practice.

There are no diagnostic magic bullets, nor are there “magic methods”. There are, however, better methods. Although much of the debate over how to predict neonatal bacterial infections has focused on which test and which cutoff, it is time that we began to debate the methods we use to measure test performance, rather than just how a given test performs. This issue has been the subject of an excellent summary by Fowlie and Schmidt (13). The use of severity of illness scores and likelihood ratios is becoming the methodologic standard in many specialties. By incorporating one of these methods (use of severity scores), Chiesa et al. (7) have taken an important step. Now it is up to the rest of us to continue further.

References