Subject Strapline: CLINICAL MICROBIOLOGY

Combining pathogen and host metagenomics for better sepsis diagnostics

Vanya Gant[1*], Mervyn Singer[2]

[1]Department of Microbiology, UCLH NHS Foundation Trust, London, United Kingdom

[2]Bloomsbury Institute of Intensive Care Medicine, University College London, London, United Kingdom.

*Corresponding author: vanyagant@nhs.net

Combining simultaneous host and pathogen metagenomic profiles in a cohort of hospitalized and critically ill patients allows for more accurate diagnosis of sepsis.
[Author: OK? one-line standfirst is easier for the reader to scan in order to make the decision to dig in to the rest of the piece].

Sepsis is defined in the clinic as an assemblage of various failing physiology and laboratory markers of organ function triggered by an infection **[AU: ok?]**. Efforts have been made to provide tighter definitions and criteria for sepsis, to achieve better and more focused clinical care, research and epidemiology[1]. The problem is that many non-infectious conditions mimic sepsis[2]. Accordingly, empiric broad spectrum antibiotic therapy is deployed to treat sepsis, even in cases with no true bacterial infection, with consequences for Global antibiotic resistance rates. For example, patients with COVID-associated sepsis are often treated with broad spectrum antibiotics to avoid missing secondary bacterial infections. In other cases, sepsis is diagnosed late, if at all, with potentially adverse outcomes[3]. Thus, clinical implementation of accurate and rapid diagnostics for sepsis at both pathogen and host levels remains challenging and urgently needed[4]. In this issue of *Nature Microbiology*, Kalantar and colleagues move one step closer to this goal and develop a sepsis diagnostic tool integrating information from both host and pathogen.

The authors used host and pathogen metagenomic next generation sequencing (mNGS) from both whole blood and plasma nucleic acids sampled from patients from two US hospitals, directly admitted to the intensive care unit (ICU) from their emergency department **[AU: of which hospital? Please clarify]** between 2010 and 2018. Patients were triaged according to whether they had 'sepsis' with either a concurrent bloodstream infection or infection identified elsewhere; 'suspected sepsis' with negative microbiological culture; critical illness for reasons other than infection; and an

'indeterminant status' reflective of clinician diagnostic uncertainty. 73 out of 92 patients adjudged to not have sepsis were given antibiotics. The authors identified several distinct patterns of host response that distinguished, with decent accuracy, infectious from non-infectious, as well as viral from bacterial causes of sepsis. Similarly encouraging data have been reported by other groups in different settings[5,6,7]. The main advance of Kalantar et al's work is that it combined host and pathogen data from plasma nucleic acid into an integrated model that considerably improved diagnostic sensitivity to 97-100% and has potential use as a rule-out test. In spite of the promising accuracy of the model to predict sepsis **[AU:OK?]**, its specificity for no-sepsis was 78%, meaning that nearly one in four of patients who do not have sepsis would be misdiagnosed as having infection. This might be due to the possibility that some of the causative microorganisms (or their nucleic acid) transiently appear in the bloodstream[8], or to unrecognised secondary infections contributing to the severity of illness[9].

The authors highlight several limitations associated with their work, including an up to 24 hours' delay between blood cultures and mNGS results, and pre-sampling administration of antibiotics impacting organism retrieval, definitely from blood culture and possibly from bacterial sequence detection. While the challenge of ensuring sufficient plasma volume to obtain adequate RNA mass seems easily solved, the additional challenge of ascribing clinical significance to detected bacterial sequences could be mitigated by combined use of host-based data. The need for external validation is also rightly emphasized. Much larger patient populations will need to be recruited to ensure not only accuracy but generalisability across ethnicities, ages, countries, immunosuppression and other potential confounders[10]. Finally, would a system trained in two US hospitals perform as well elsewhere? Resolving these questions will likely be needed for registration requirements.

Although the work by Kalantar and colleagues is a valuable contribution to diagnostics research, further challenges from a clinical perspective arise. Will enough differentiation be found within the host signature to unravel systemic inflammatory conditions with clinical phenotypes more closely mimicking sepsis, such as major surgery, trauma and pancreatitis? Would the presented approach yield (near-) comparable data in pre-symptomatic or early organ dysfunction patients allowing pre-emptive intervention? Would there be added value from sequential monitoring? From a clinician's standpoint, is rapid knowledge of the etiologic pathogen(s) critical for optimal treatment of sepsis, as claimed? Arguably, a swift broader speciation (Gram positive, Gram negative, fungal, viral) with anti-infective drug susceptibilities may suffice for directing the most important drug interventions for better outcomes.

The authors have nicely demonstrated how integration of pathogen and host response analysis improves diagnostic accuracy. Further benefit might be achieved from combining clinical, biochemical, molecular, imaging and/or various -omic profiles into the mix [11]. Applying machine learning algorithms to such increasingly complex datasets would be relatively easy as Cloud Computing capabilities are accessible, powerful and cheap. The ultimate goal would be to rapidly direct effective clinical interventions for better outcomes, simultaneously targeting pathogen elimination and "personalised" therapy to the highly heterogenous host response to sepsis. It is important to keep in mind that to truly impact patient outcomes, such interventions must translate into real life actions for pragmatic solutions. **[AU: OK? Deleted as too many references]**.

Studies such as Kalantar's and colleagues provide further encouragement and direction for leveraging large clinical and laboratory datasets with computing power and algorithms. However, and some 26 years after computers could predict ICU survival[12,] with no tangible clinical uptake, it remains no easy task to leverage combined host and pathogen diagnostics at pace and scale with built-in around-the-clock capability and user-friendly solutions, at affordable cost and in clinically actionable timeframes. The challenge to develop a road map for the use of genome-based diagnostics beyond academic proof-of-concept for better patient treatment should now be the priority.

Competing interests

In the last three years Singer has undertaken research projects and/or sat on advisory boards for a number of sepsis and infection diagnostic companies including Abbott, Biomerieux, deepUll, Defence Science and Technology Laboratory (dstl), Gentian, Roche Diagnostics, Safeguard Biosystems

In the last three years Gant has undertaken research projects and/or sat on advisory boards including BIomérieux, Oxford Nanopore, Curetis, Vidya

References

1. Singer M *et al*. *JAMA315, 801-810 (*2016)
2. Klein Klouwenberg *et al*. Crit Care; 19:319. https://doi.org/10.1186/s13054-015-1035-1 (2015)
3. Filbin, MR *et al. Crit Care Med 46, 1592–9 (2018)*
4. *Enne VI et al.* Thorax Jan 2022, thoraxjnl-2021-216990; DOI: 10.1136/thoraxjnl-2021-216990

5. Sweeney TE, Wong HR, Khatri P.. *Sci Transl Med.; 8:346ra91 (2016)*
6. Bhattacharya S, *et* al. *Sci Rep.; 7:6548 (2017)*
7. Gómez-Carballa A, *et al. Int J Mol Sci.; 22:3148- (2021)*
8. Bahrani-Mougeot FK *et al, J Clin Microbiol 46; 2129 -2132 (2008)*
9. Tan *et al*., Microbial Genomics;7:000642, (2021)
10. Bodkin N *et al. Genome Med*; 14:18 (2022)
11. Maslove DM *et al*. *Nat Med 2; 1141–8 (2022)*
12. Dybowski R *et al. Lancet* 347:1146-1150 (1996)

.

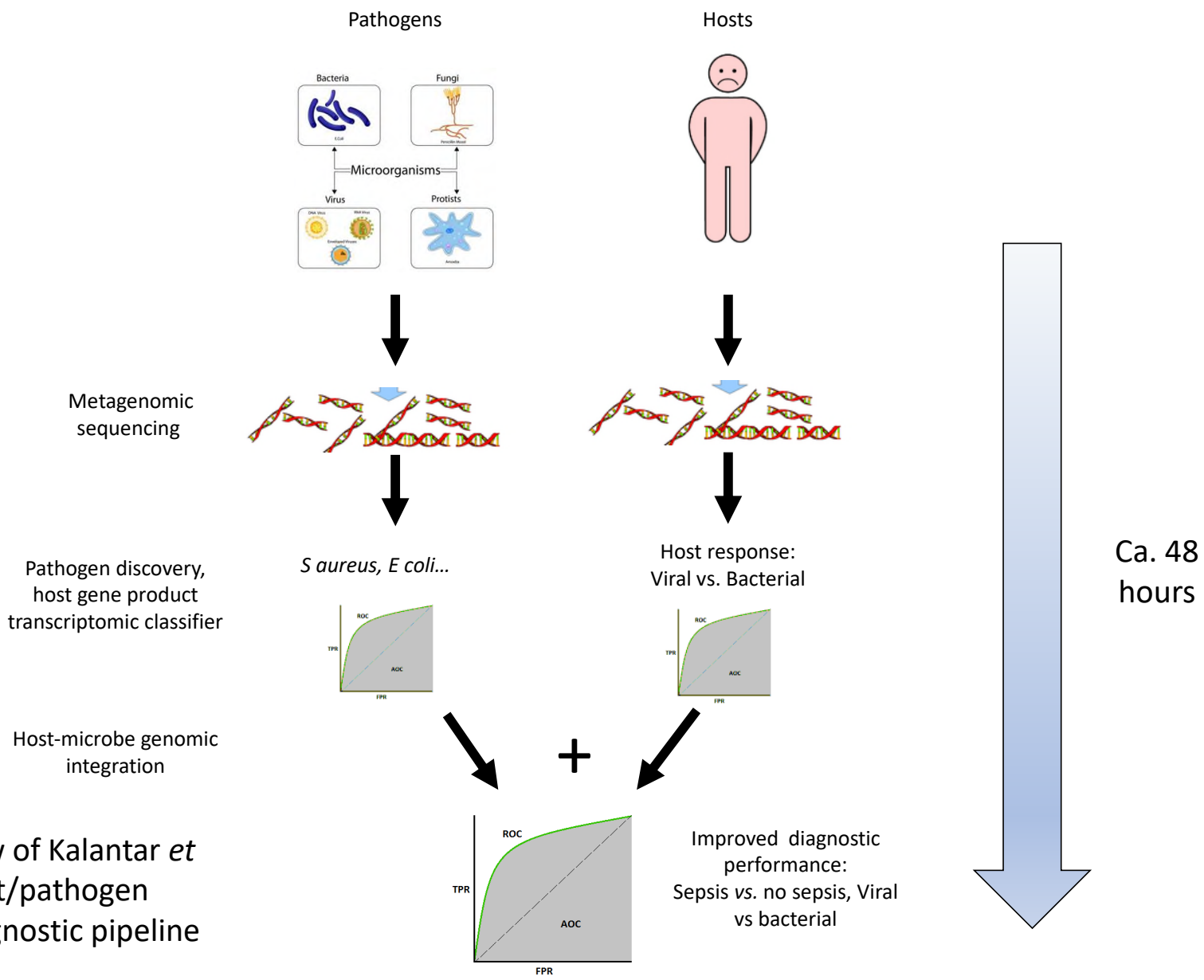

Figure 1. Overview of Kalantar *et al's* combined host/pathogen metagenomic diagnostic pipeline