Commentary on a combined approach to the problem of developing biomarkers for the prediction of spontaneous preterm labor that leads to preterm birth

R.F. Lamont a,b,*, L.S. Richardson c, J.J. Boniface d, T. Cobo e, M.M. Exner f, I.B. Christensen g, S. K. Forslund h, A. Gaba i, H. Helmer j, J.S. Jørgensen a,k, R.N. Khan i, T.F. McElrath m, K. Petro f, M. Rasmussen n, R. Singh s, R.M. Tribe o, J.S. Vink p, C.A. Vinter a, N. Zhong q, R. Menon c,

a Research Unit of Gynaecology and Obstetrics, Department of Gynaecology and Obstetrics, Odense University Hospital, Institute of Clinical Research, University of Southern Denmark, Odense, Denmark
b Division of Surgery, Northwick Park Institute for Medical Research Campus, University College London, London, UK
c Dept of Obstetrics and Gynecology, The University of Texas Medical Branch at Galveston, Dept. Electrical and Computer Engineering Texas A&M University, College Station, TX, USA
d Sera Prognostics, Inc., 2749 East Parleys Way, Suite 200, Salt Lake City, UT, 84109, USA
e BCNatal - Barcelona Center for Maternal-Feal and Neonatal Medicine (Hospital Clinic and Hospital Sant Joan de Deu), Institut Cliní del Ginecología, Obstetricia i Neonatologia, Fetal i D Fetal Medicine Research Center, Institut d Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona. Barcelona, Spain
f Centre for Biomedical Research on Rare Diseases (CIBER-BBN), Barcelona, Spain
g Hologic, Inc., 10210 Genetic Center Dr, San Diego, CA, 92121, USA
h ARGEND Biotec ApS, Aarhus, Denmark
i Experimental and Clinical Research Center, A Cooperation of Charité-Universitätsmedizin and the Max-Delbrück Center, Berlin, Germany
j Department of Obstetrics and Maternal-fetal Medicine, Vienna Medical University, Austria
k Centre for Innovative Medical Technologies (CIMT), Odense University Hospital, Klovvervænget 8, 5000, Odense C, Denmark
l Odense Patient Data Explorative Network (OPEN), Odense University Hospital/University of Southern Denmark, J. B. Winlows Vej 9 a, 3. Floor, 5000, Odense C, Denmark
m Division of Medical Science and Graduate Entry Medicine, School of Medicine, University of Nottingham, Room 4115, Medical School, Royal Derby Hospital Centre, Derby, DE22 3DT, UK
n Brigham & Women’s Hospital, Boston, USA
o MRVIE Inc., 820 Dubuque Ave., South San Francisco, CA, 94080, USA
p Dept. of Obstetrics and Children’s Health, School of Life Course Sciences, King’s College London, St Thomas’ Hospital Campus, London, SE1 7EH, UK
q Division of Maternal Fetal Medicine, Department of Obstetrics and Gynecology, Columbia University Medical Center, New York, NY, USA
r New York State Institute for Basic Research in Developmental Disabilities, 105 Forest Hill Road, Staten Island, NY, 10314, USA

Introduction: Globally, preterm birth has replaced congenital malformation as the major cause of perinatal mortality and morbidity. The reduced rate of congenital malformation was not achieved through a single biological or biochemical marker at a specific gestational age, but rather through a combination of clinical, biological and biochemical markers at different gestational ages. Since the aetiology of spontaneous preterm birth is also multifactorial, it is unlikely that a single biomarker test, at a specific gestational age will emerge as the definitive predictive test.

Methods: The Biomarkers Group of PREBIC, comprising clinicians, basic scientists and other experts in the field, with a particular interest in preterm birth have produced this commentary with short, medium and long-term aims: i) to alert clinicians to the advances that are being made in the prediction of spontaneous preterm birth; ii) to encourage clinicians and scientists to continue their efforts in this field, and not to be disheartened or
nihilistic because of a perceived lack of progress and iii) to enable development of novel interventions that can reduce the mortality and morbidity associated with preterm birth.

Results: Using language that we hope is clear to practising clinicians, we have identified 11 Sections in which there exists the potential, feasibility and capability of technologies for candidate biomarkers in the prediction of spontaneous preterm birth and how current limitations to this research might be circumvented.

Discussion: The combination of biophysical, biochemical, immunological, microbiological, fetal cell, exosomal, or cell free RNA at different gestational ages, integrated as part of a multivariable predictor model may be necessary to advance our attempts to predict sPTL and PTB. This will require systems biological data using “omics” data and artificial intelligence/machine learning to manage the data appropriately. The ultimate goal is to reduce the mortality and morbidity associated with preterm birth.

1. Introduction

Five decades ago, congenital malformation was the major cause of perinatal mortality and morbidity in high-income countries [1]. With developments in obstetric ultrasound and biochemical markers, congenital malformation was replaced by preterm birth (PTB) as the major cause of perinatal mortality and morbidity, initially in high-income countries but now globally, in high, medium and low-income countries. The financial cost of PTB is also an enormous burden on healthcare resources, and the psychosocial effects of PTB are immeasurable [2,3]. The reduction in the rate of congenital malformation has been achieved, not through a single biophysical or single biochemical marker at a specific gestational age, but rather a combination of clinical, biophysical and biochemical markers at different gestational ages. These markers are then linked through complex algorithms or software to come up with a measure of risk, which leads to further diagnostic tests and eventually novel interventions. Spontaneous preterm labour (sPTL) leading to PTB is a heterogeneous condition, with a multifactorial aetiology. However, while the pathophysiology is different, the final end common pathway of progressive uterine contractility, cervical changes and membrane rupture are the same, and hence is often referred to as the “Preterm Parturition Syndrome” [4]. Since the aetiology of sPTL and PTB is multifactorial [5], associated inter alia with a genetic predisposition, inflammation/infection, allergy, uterine stretch, vascular disorders and stress, it is unlikely that a single biomarker test at a specific gestational age will emerge as the definitive predictive test. Accordingly, it will be necessary to explore different biomarkers at different gestational ages and to combine these in a way that, in combination, rather than individually, they may be more beneficial for the prediction of sPTL leading to PTB.

The Biomarkers Group of PREBIC (The Preterm Birth International Collaborative), (http://www.prebicglobal.org), which comprises clinicians and basic scientists with a particular interest in PTB, have put together this commentary with short, medium and long-term aims. In the short term we aim to alert clinicians to the advances that are being made in the domain of predicting sPTL leading to PTB, and we plan to do this using clear language that is comprehensible to practising clinicians. In the medium term, we hope this will encourage clinicians and scientists to continue their efforts in this field and not to be disheartened or nihilistic because of a perceived lack of progress. We also hope that this might not only result in a practical, targeted and effective means of identifying women at high risk of sPTL leading to PTB but might provide additional information about the mechanisms and pathways involved in the process. In the long term, we aim to enable development of new interventions that can reduce the mortality and morbidity associated with PTB.

Each section provides a brief overview by clinicians, basic scientists and other experts in the field, to highlight and prioritise the potential, feasibility and capability of these technologies in the prediction of sPTL with subsequent PTB. We outline recent advances in candidate biomarkers for the prediction of sPTL leading to PTB and discuss how current limitations to this research might be circumvented.

1.1. Circulating microparticles

Circulating microparticles (CMPs) are nanosized extracellular vesicles that are composed of a lipid bilayer [6]. Although modern science only discovered CMPs in the late 1980s, they have been a feature of eukaryotic physiology for centuries [7–9]. In human physiology, CMPs have been isolated from a diverse variety of fluid media including saliva, amniotic fluid, ascites, semen, cerebrospinal fluid and, perhaps most importantly for the present commentary, plasma [6]. CMPs are produced by a wide variety of cell types including platelets, endothelial cells, erythrocytes, trophoblastic cells, lymphocytes and almost all malignancies so far investigated [6,9,10]. Functionally, CMPs represent a “liquid biopsy” that can be obtained in physiological “real time” [11,12]. CMPs can be divided into three categories: i) exosomes (50–150 nm); ii) microvesicles (100nm-1μm) and iii) apoptotic bodies (200 nm–5 μm) [20]. CMP contents, including mRNA, microRNA, DNA fragments, as well as both cytosolic and surface proteins [13,21,22], are intended to be functional mediators of both proximal and distal cellular function [21]. Proteins can also be electrostatically associated with the surface of the CMP bilayer, although the functional significance of these proteins is less clear.

CMP signaling is highly active in pregnancy, and in the murine model, CMP effects can modify maternal physiology at all gestational ages with specific modulation of systemic inflammation and intrauterine tolerance [13,14]. In humans, CMPs can be detected in maternal circulation as early as six weeks post-menstrually with an exponential increase in concentration as pregnancy progresses [10,15]. Maternal plasma has been sampled between 10 and 12 weeks of gestation from which distinct CMP associated protein patterns in women who deliver spontaneously before 35 weeks of gestation can be identified [16–18]. Other groups have identified CMP patterns that are associated with the subsequent diagnosis of preeclampsia [19–24]. This suggests that CMP associated signaling represents a potent form of “early warning” or risk stratification for adverse pregnancy outcome.

The clinical use of CMP signaling in perinatal medicine is challenging. There are many techniques to isolate CMPs including differential centrifugation, density gradient centrifugation, ultrafiltration, precipitation, and size exclusion chromatography [25]. Most of these are relatively labor-intensive and time-intensive and may not be ideal for scaled-up commercial clinical application. Furthermore, reproducibility is a problem, so attempts have been made to resolve this issue, but larger scale validations will be required [18].

To date, the most widely studied medium for CMP collection has been plasma. This indicates that what is being studied is circulating, systemic CMPs, and not necessarily placental or trophoblast specific particles. However, to develop a prognostic test, this distinction is less important. Whether other media such as saliva or cervico-vaginal secretions might provide a more predictive yield of CMPs has not been widely investigated. Regardless of the media from which CMPs are collected, they are remarkably durable. Unlike free protein or nucleotides, CMP associated proteins retain a discernable signal even after days of storage at room temperature. In addition, CMPs have the potential to question the physiologic state of specific tissues at the time of their release, which could prove to be very informative in perinatal diagnosis.
Finally, one of the most exciting aspects of this area of research is the potential to use CMP protein analysis to classify molecularly the various phenotypes within a larger category of perinatal pathology. The suggestion that preeclampsia, like sPTL and PTB, is a syndrome rather than a single pathology is becoming widely accepted [26]. If this is the case, then expecting a single screening test or prophylactic measure to predict or treat the syndrome is optimistic. However, the ability to test for, or treat specific phenotypes may be more sensitive and specific. CMPs are well suited for this level of discernment since, by their nature, they reflect the specific mechanisms or pathways involved in the disease process. Targeted testing and treatment would allow a more accurate diagnostic and clinical assessment.

1.2. Cell-free RNA

As cells replicate in the growing fetus, RNA from expressed genes is released from the placenta into the maternal circulation. Although RNA is known to be short-lived, several studies [27-29] have demonstrated that RNA can be readily found in maternal plasma, raising the possibility of a non-invasive test to monitor the fetoplacental development by following changes in cell free RNA (cfRNA). cfRNA provides a snapshot of information about which fetomaternal genes are expressed and at what levels. By comparing expression patterns between healthy and diseased states, sets of biomarkers can be identified that may advance our knowledge with respect to the underlying causes, by revealing affected pathways. Promising results for the prediction of sPTL in a high-risk population using a limited set of cfRNA markers has been obtained [27]. Pregnant women with a history of previous PTB were sampled between 23 and 28 weeks of gestation. RNA was extracted from plasma and sequenced. Using machine-learning algorithms, a set of predictive markers was identified, and these were validated in a qPCR assay in a separate group of high-risk women, as well as an independent group of 31 term pregnancies. This resulted in an area under the curve (AUC) of a receiver operating characteristic (ROC) curve of 0.81 in the validation cohort. With multiple aetiologies for sPTL leading to PTB, it is important to explore large and diverse sets of samples. This will allow the identification of subsets of women with different risk categories and identify specific signatures within each group. In summary, cfRNA is a promising avenue for new diagnostics and hypothesis building.

1.3. Proteomics

1.3.1. Proteomics approaches and applications in PTB research

Proteomic analysis involves the measurement of proteins and their function in the expression of an organism’s genome in health and disease.Advantages of protein-based biomarkers over other biomolecules generally relate to their close association with phenotype and relative stability. Since the study of proteomics began in the mid-1990s [30,31], technology has greatly improved to involve fluorescent labelling in 2-D gel electrophoresis [32], the development of protein arrays [33], antibody [34] and nucleic acid [35] affinity reagent-based multiplexed measures of proteins, and evolving mass spectrometry (MS) based strategies [36]. Proteomics of extracellular vesicles including exosomes is covered in Section 1 above. Here, we focus on the application of MS approaches to protein biomarker discovery in PTB. To develop PTB biomarkers, we need to understand how the biochemical pathways underlying normal pregnancy are disrupted in abnormal pregnancies. During discovery, shotgun proteomic approaches that measured hundreds of proteins simultaneously, were typically employed, although throughput was limited [37]. Longitudinal analyses were carried out to measure how protein levels change as pregnancy progresses [38,39]. Lysates from placental tissues and explants and other biofluids such as cord blood, amniotic fluid, maternal serum/plasma, cervico vaginal secretions, urine, and saliva have been used to compare proteomes from these delivering term versus preterm or with other pregnancy complications such as preeclampsia, typically from single time points [40]. While informative, these analyses have not lead to validated biomarkers due to the small sample numbers, sampling issues such as timing of collection and demographics, and heterogeneity in the aetiology of PTB [40]. More recently, applying a systems biology strategy to the proteomic analysis of a large cohort and three independent phases of development, resulted in a clinically validated predictor of sPTL and PTB [41].

1.3.2. The future of proteomics

All biomarker detection platforms have limitations in multiplexing or dynamic range [42,43] that require some sample fractionation or enrichment to measure the wide array of biomarkers of interest. Such fractionation may occur pre-detection, as with immune-affinity capture methods, or online with chromatographic or ion mobility separations [44-49]. Notwithstanding such approaches, the increasing sensitivity of various platforms and techniques such as MS enable the measurement of analytics that have not previously been possible to detect. Regardless of the technology used, future biomarker assays will require ever increasing automation and miniaturization to deliver the capacity, throughputs and economics to enable widespread distribution and adoption.

1.4. Fetal cells in maternal blood as Biomarkers for the prediction of PTB

The placenta and fetoplacental membranes and tissues of fetal origin, play important roles in the maintenance of pregnancy and the onset of term and preterm parturition [50,51]. This supports the hypothesis that they may contain useful predictive biomarkers to determine fetomaternal pathophysiology. Recently, cell based non-invasive prenatal testing (cbNIPT) approaches using fetal cells isolated from maternal blood to diagnose fetal abnormalities has provided unique information. An example is full fetal genome versus selected chromosomes. Comparison is more sensitive than other diagnostic testing derived from fetal samples isolated from maternal blood because cell free DNA. A novel adaptation of the cbNIPT approach, using maternal blood to target cells derived from the placenta and connecting membranes, could offer a new approach to the prediction of PTB. In cbNIPT, cells are extracted from maternal blood using a cocktail of fetal cell specific antibodies and processed for analysis of the fetal genome [52,53]. By using this technique, researchers may be able to extract cells from fetal membranes circulating in maternal blood and use them as biomarkers to predict PTB. In cases of term labour, sPTL and preterm labor rupture of membranes (PPROM), microfractures, or areas of membrane remodelling and cellular turnover in the fetal membranes has been observed [54,55]. These findings have led to the hypothesis that cells passing through these microfractures might be shed into maternal blood and could be used as biomarkers of sPTL [56]. Pilot studies using RNA-sequencing of fetal membrane cells from the amnion and chorion and maternal blood cells, have identified genes that are over-expressed in the former compared to the latter. In a pilot study, a few markers were tested and were found capable of enriching and identifying fetal membrane cells from maternal blood at >35 weeks of gestation (unpublished data). In future, it is hoped that isolated cells from maternal blood could be used to predict sPTL and PTB by studying the quantity, expression, and phenotype of these cells at different gestational ages. Researchers need to confirm the specificity of isolated cells to identify fetal, placental, and fetal membrane derivatives for individual analysis. This approach has clinical potential and may provide a novel biomarker for sPTL and PTB.

1.5. Fetal fibronectin

In current clinical use, fetal Fibronectin (fFN) is the most established biochemical biomarker of cervicovaginal fluid that provides information on the risk of PTB. Qualitative and quantitative fFN has proven useful for the prediction of PTB in both asymptomatic and symptomatic women [57-63]. Other options such as Partasure® (http://partasure.com/) and
Actim Partus® (https://www.abbott.com/) for symptomatic women, while cheaper, have yet to achieve widespread clinical use due to limited evidence.

Discovered in the mid 1980s [64,65], fFN is an extracellular matrix glycoprotein produced by amniocytes and cytotrophoblast [60]. It is thought to be present mainly at the choriodecidual interface, a union between maternal and fetal tissues. Mechanistically, this is an area that is poorly studied, but there is evidence that inflammatory pathways cause mobilisation of fFN [66,67].

Highlighted as a potential clinical tool [60], a qualitative membrane immunosorbert ELISA assay [ Rapid fFN for the TLiQ® System and QuickCheck fFN (https://www.hologic.com/) ] was developed. The qualitative test has been FDA approved for use as an aid to evaluate rapidly the risk of PTB in women with signs and symptoms of sPTL. This has been developed into a quantitative test for use in Europe and elsewhere (Rapid fFN 10Q Cassette Kit) to stratify the risk of PTB associated with increasing concentrations of fFN [58]. For asymptomatic women at risk of PTB, the recommended gestational age range of application is 22–27 completed weeks of gestation, but it may also be useful from 18 completed weeks of gestation and beyond [62].

The advantage of the qualitative fFN test is that it has been proven to assist and guide clinical management such as the use of in utero transfer or administration of antepartum glucocorticoids [68,69]. Quantitative fFN test results can be incorporated, with or without cervical length measurement, into prediction models for the risk of PTB. Published algorithms have proved useful for the management of pregnant women with symptoms suggestive of abnormal or preterm uterine activity (23–34 completed weeks of gestation) and high-risk asymptomatic women (18–36 completed weeks of gestation) [70,71]. Algorithms for asymptomatic women with singletons and twins can be used but women should have at least one of the following risk factors: i) previous cervical surgery; ii) previous PPROM; iii) previous PTB (24–36 completed weeks of gestation) and iv) previous late miscarriage (16–23 completed weeks of gestation). The algorithms are accessible for clinicians through the Quipp website (https://quipp.org) or the free mobile/cell phone device Quipp app [71,72]. The impact of using the app in UK obstetric units is being evaluated in a large cluster trial [73].

However, quantitative fFN is best used for prediction from 18 to 20 weeks of pregnancy onwards so there is still a need to seek tests that could be used earlier in pregnancy. The current cost of fFN tests and relevance in low-income countries where the burden of PTB is highest needs to be addressed. The technology may be improved by combining fFN testing with other biomarkers such as transvaginal ultrasound of cervical length and molecular microbiological testing of vaginal eubiosis or dysbiosis (see Section 6 below).

1.6. The relationship Between the eubiotic or dysbiotic vaginal Microbiome for the prediction of PTB

1.6.1. Eubiotic or dysbiotic vaginal microbiota using culture-independent techniques

Lactobacilli provide vaginal eubiosis through numerical dominance, vaginal epithelial adhesion, and killing ability over dysbiotic organisms through production of lactic acid, H₂O₂ (hydrogen peroxide), and naturally occurring antimicrobials like bacteriocins. Culture only identifies Lactobacilli to the genus level, so beneficial species-specific function cannot be determined [74]. Vaginal dysbiosis, mainly in the form of bacterial vaginoses (BV), is strongly associated with PTB, particularly when detected early in pregnancy [75]. Unfortunately, culture cannot diagnose BV and currently the gold standard for diagnosis is quantitative Gram stain microscopy, which, though more objective that wet mount microscopy, is still somewhat subjective. With new information from cultivation-independent techniques [76], we can now identify the species-specific functions of Lactobacilli and we now know that worldwide, vaginal eubiosis comprises one or two species of Lactobacilli from a shortlist of four: L. crispatus, L. gasseri, L. iners, and L. jensenii, corresponding to community state types (CST) I, II, III, and V, respectively [77], similar to the findings of “enterotypes” clustering in the human gut microbiota [78]. These techniques have also identified BV associated bacteria which were either previously unidentified or underappreciated with respect to BV, and dominate CSTs IVa and IVb [77].

1.6.2. Current evidence

Currently, the evidence for the use of vaginal microbiome studies to predict PTB is sparse, less than robust, heterogeneous, conflicting, and of poor quality, such that a minimum database of recommendations for future studies has been proposed by PREBIC [79]. Candidate organisms associated with vaginal dysbiosis in the form of BV for molecular-based techniques include: i) BV associated bacterium (BVAB)-1, –2 and –3; ii) Gardnerella vaginalis; iii) Atopobium vaginae; iv) Megasphaera spp (Types –1 or –2) and v) Lactobacillus/Sneathia species. A growing number of studies link these species, either individually or synergistically to an elevated risk of PTB. However, there is heterogeneity with respect to racial groups, and the extent to which this reflects genetic as opposed to sociocultural factors is unclear.

1.6.3. Commercially available molecular diagnostic tests

New technologies employing molecular markers for BV have been developed to overcome the problems associated with Gram stain microscopy diagnosis and point of contact tests. Molecular based tests have the advantage of: i) objectivity; ii) quantification; iii) detection of fastidious organisms and iv) validity for self-obtained vaginal swabs. The techniques are dependent upon the detection of specific bacterial nucleic acids and are primarily available as direct probe assays [80] or nucleic acid amplification test (NAAT) assays. Though there are many commercially available laboratory developed tests, to our knowledge, currently in the European Union and USA, there are six commonly used, molecular diagnostic tests for the diagnosis of BV [81–84], all of which are multiplex PCR assays, two of which are approved by the FDA [82, 83], and three CE-IVD (Conformité Européenne in vitro Diagnostic) approved [82–84]. These have been reviewed elsewhere [85].

Future research into vaginal dysbiosis in the form of BV for the prediction of PTB may need to repeat much of the BV research based on Gram stain microscopy with cultivation-independent molecular based techniques. This might also address the complex confounding factors that independently affect the risk of PTB and the effect of the vaginal microbiota. Such molecular techniques have major advantages over Gram stain microscopy with respect to continuation, but those currently available commercially have their limitations. Not all consider the absence of organisms associated with eubiosis at the same time as the presence of organisms associated with dysbiosis as targets for the diagnosis of BV for the prediction of PTB. By necessity, they target a small number of organisms that may not reflect less common subtypes of the BV syndrome, which may have different aetiologies, different culture profiles, different response to antibiotics and different phenotypic outcomes. Understandably, current commercially available molecular based diagnostic tests do not include L. iners as a target organism because it does not differentiate between BV and eubiosis, yet L. iners is likely to prove to be a very important organism in future use of the vaginal microbiome as a predictor of sPTL and PTB, [79,86,87].

1.6.4. Requirements to improve technology/techniques

Recommendations for a minimum database for future studies on the relationship between the vaginal microbiome and PTB should be followed [79]. When conducting molecular-based techniques in relation to PTB, consideration should be given to the clades/biovars/V-regions of primers for organisms such as G. vaginalis and U. urealyticum [79,88–90]. Investigation of the vaginal virome and mycobiome should be pursued further, as should the extent to which the gut microbiome may act as a reservoir for the vaginal microbiome. This may provide additional, special and temporal information over a longer time-scale that is different and potentially more relevant with respect to the short-term
fluctuations that occur in the vagina. The shift that occurs throughout pregnancy and how this differs by race should also be considered [91]. Finally, while four CSTs have been identified that relate to a Lactobacillus spp dominated vaginal microbiota (I, II, III and V) and two dysbiotic CSTs (IVA & IV b) [77], the possibility of a number of as yet undefined or unrecognised additional dysbiotic vaginal CSTs that may have different aetiologies (including sexually transmitted sub-types), different culture based microbiological profiles, different responses to antimicrobial therapies, and different phenotypic outcomes such as sPTL and PTB should be considered. Such additional CSTs may not be represented in the choice of targeted organisms in current commercially available molecular based diagnostic tests [81–84].

1.6.5. Future research

It is not sufficient to know the composition of the vaginal microbiota, without knowing: i) the metabolome of the vaginal milieu created by a given microbiome; ii) the host response to the combination of the microbiome, and iii) the metabolome it generates. We now understand better the interaction and balance between “exposure” and “susceptibility”, and this can be considered a gene-environmental interaction [92]. Balance is important when considering cytokine profiles at any given body site. The cytokine response should consider not only the pro-inflammatory but also the anti-inflammatory cytokines together with the host tolerance to a particular microbiota and the subsequent outcome.

Furthermore, we must emphasise that sPTL which leads to PTB is a syndrome with different aetiologies, different phenotypic presentations, different management options, and different phenotypic outcomes. These different phenotypic outcomes should be viewed separately when considering the aetiology, prediction and prevention of PTB. Depending on the combination of influences from the various vaginal microbiome components, the local milieu created and the host response, the phenotypic outcome may range from normal term birth to PPROM to sPTL with intact membranes (with or without vaginal bleeding) to preterm stillbirth and other phenotypic outcomes [93]. Tracking the range of these outcomes using data collected from multiple systems will allow better resolution of their shared and divergent mechanisms.

1.7. Immunological markers

Regardless of whether or not there is detectable microbial invasion of the amniotic cavity (MIAC), intra-amniotic inflammation (IAI) is considered to be an essential component of sPTL and PTB at early gestations. Approximately 40% of women in sPTL before 28 completed weeks of gestation will have MIAC and/or IAI [94]. IAI is associated with an adverse neonatal outcome [95] and there is a clear clinical need to diagnose these inflammatory/infectious condition for the management of a woman admitted with threatened sPTL at early gestations. The main barrier to the diagnosis of infection/inflammation is the requirement to perform an invasive procedure such as amniocentesis. In addition, there is no clear definition of IAI. Concentrations of interleukin (IL)-6 have been identified as a useful biomarker but consensus about a clinically relevant cut-off is required [94,96–98] to ensure reproducibility and external validation.

1.8. Inflammation and whether infective or sterile

It is important to differentiate between infection or inflammation, particularly microbial associated intra amniotic inflammation, sterile inflammation and MIAC. Several authors have proposed different biomarkers of MIAC and/or IAI in amniotic fluid [99–102], using minimally invasive biological samples such as maternal blood [104], or cervico-vaginal secretions [103,105–108] [109], using hypothesis-based biomarker discovery approach such as ELISA or “omics” studies. However, none of these markers have yet to show sufficient accuracy to be used as stand-alone predictors in clinical practice.

Multivariate predictor models can improve diagnostic performance of single biomarkers alone. In women with symptoms of sPTL, multivariable predictors models of MIAC and IAI have been reported with good diagnostic performance [110,111]. However, these have not been properly validated [111] or have not been developed sufficiently [110] to be used as a tool for rapid diagnosis in a clinical setting.

Currently, clinicians are still unable to treat MIAC and/or IAI, but an early diagnosis of MIAC and/or IAI would assist in the planning of the “expected” delivery of these babies and management using antenatal strategies that have shown a perinatal benefit such as steroids [69], neuroprotective agents such as magnesium sulfate [112] or antibiotics [113]. Conversely, a good biomarker would avoid unnecessary over-treatments in the low-risk group of women who are unlikely to have a PTB in the following 7 days (the “rule out strategy”). Since most clinicians and patients are reluctant to perform an amniocentesis, there is a need to develop clinically feasible minimally invasive tests that target and identify accurately the high-risk group of MIAC and/or IAI who might benefit from the collection of amniotic fluid. This should reduce the number of amniocenteses for this indication.

1.9. Ultrasound biomarkers to predict premature cervical remodelling and spontaneous preterm birth

1.9.1. Methodology/clinical practice

Ultrasound-based techniques to detect premature cervical remodelling have been categorized as methods that either detect cervical tissue deformation such as cervical length measurement, cervical consistency index and static strain elastography or attempt to detect or quantify changes in cervical tissue extracellular matrix such as shear wave elastography, and backscatter power difference, attenuation [114–121]. For details of these techniques, the reader is referred to comprehensive reviews published elsewhere [120,121]. To date, although these techniques may hold promise, cervical length measurement is the only biophysical method that is commonly used in the clinical setting to identify women at risk for sPTL leading to PTB. Currently, various recommendations exist with respect to if, and when, to perform cervical length screening. The Society for Maternal Fetal Medicine recommends routine transvaginal ultrasound cervical length screening between 16 and 24 completed weeks of gestation for women carrying a singleton pregnancy and with a history of a prior sPTL and PTB [122]. While guidelines and recommendations differ from country to country, the American College of Obstetricians and Gynecologists (ACOG) does not recommend routine cervical length screening [123]. Rather, the ACOG recommends that the cervix should be examined when technically feasible [124]. While it is controversial to recommend universal cervical length measurement in asymptomatic women without a previous PTB, the International Federation of Gynecology and Obstetrics (FIGO) recommends performing cervical length screening in all women between 19 and 23 completed weeks of gestation [125].

1.9.2. Advantages/disadvantages

Although cervical length screening has been shown to be beneficial in identifying women with a short cervix, only ~30% of women with a short cervix, who do not undergo an intervention, eventually deliver before 35 completed weeks of gestation [126]. In addition, cost-effectiveness analyses models have not shown consistently that universal cervical length screening is cost-effective [127]. Finally, although the techniques to measure cervical tissue deformation, or changes in cervical tissue extracellular matrix have shown promise, technical challenges still exist with each technique [120,121], and larger studies are needed to confirm their results, and their role as biophysical markers to predict sPTL leading to PTB. Another disadvantage is the cost of equipment together with the need for trained personnel. In addition, there are concerns as to how such techniques can be standardised worldwide.
1.9.3. Requirements to improve technology/technique

A limitation to current ultrasound-based biomarkers is that they only evaluate the cervix. Recent studies using ultrasound-based, anatomically correct computer simulation models of the pregnant pelvis suggest that cervical function in pregnancy is influenced not only by cervical tissue structure, but also its surrounding anatomical and mechanical environment [128,129]. Given the complexity of pregnancy and the various pathophysiology/phenotypes of sPTL and PTB, it is likely that future biomarkers should not rely on a single aspect of cervical tissue but will need to take a more holistic view of a pregnant woman’s anatomy. An ultrasound-based, personalized computer simulation model that incorporates a woman’s specific anatomical measurements, reproductive tissue mechanical and biochemical properties and pregnancy information, poses a comprehensive and exciting possibility as a biomarker to predict sPTL leading to PTB. Accordingly, it may be possible to construct an algorithm to include a number of biophysical and biochemical biomarkers.

1.10. Systems biology: an integrated approach to identify and characterize biomarkers

Systems biology has been defined as the study of dynamic interactions within biological networks, built on the increasing availability of high-throughput data to which increasing computational capacity can be applied. These interactions can give rise to emergent properties unpredicted by any of the components. In this sense, systems biology can be viewed as the integration of multiple biological data across different levels of structure and scale [87]. Applying systems biology to the identification of biomarkers of PTB, integration is performed on multiple levels.

The first level is to integrate clinical information, which may provide a base for biological assessment. Such clinical information should include, but is not limited to, the history of pregnancy, the pathophysiological condition of pregnant women and family member(s), the prenatal care received, and clinical manifestations as well as clinical management. In addition, the clinical information may need to include clinical laboratory generated data, which must be standardized to avoid batch effects or other technical artifacts.

The second level is to integrate various “omics” data, which include data generated from DNA sequencing, that gives genomic variants, including DNA single nucleotide variations (SNVs). These comprise genomics, transcriptomics and epigenomics: i) DNA variants such as SNP, CNV or InDel; ii) epigenetics variants such as methylation status, chromatin remodeling and histone modifications; iii) RNA variation such as differential expression, non-coding RNAs; iv) proteomics; v) metabolomics and vi) metagenomics.

In addition to integration of clinical or “omics” data in one dimension, systems biology must perform multidimensional integrations at the third level. During pregnancy, longitudinal data may be collected or generated at a number of time periods. By integrating these time-point data longitudinally, the dynamics of the biomarker can be followed through to the outcome of pregnancy. This is important, since features that may be observed at a certain gestational age in a subsequent term pregnancy may be an indicator of high risk of PTB when observed at an earlier or later gestational age.

Systems biology multi-dimensional assessment of data will provide a powerful platform to analyze and record the pathophysiological pathway of pregnancy, particularly adverse outcomes of pregnancy such as PTB [88,89]. Applying systems biology to the search for biomarkers may maximize reproducibility while minimizing bias or errors that often occur in one-dimensional or a single “omics” approaches.

1.11. Bioinformatics, computational statistics, artificial Intelligence and machine learning in biomarker discovery, Analysis and selection

The Great Obstetric Syndromes [121] comprising PTB, preeclampsia, IUGR and intrauterine fetal death, together with metabolic disease in the form of gestational diabetes mellitus and obesity, contribute significantly to fetomaternal and neonatal morbidity and mortality, and disease in later life for both mother and child [122]. The links across these conditions may include insulin resistance, immune modulation and inflammation/infection, oxidative stress, and the host’s response to these. The aetiology of each is multifactorial and the evidence with respect to aetiology, prediction, prevention and intervention from conventional statistical techniques is less than robust, confusing, and often counterintuitive.

It should be possible to employ novel bioinformatics and computational statistics using artificial intelligence or machine learning applied to existing data, to identify trends that would not otherwise be evident using conventional statistical methods. By eliciting subtle changes or differences using “big data”, while carefully accounting for biases in this data, it might be possible to identify nuances in studies that identify patterns of aetiology, and to develop novel predictive, preventative and/or interventional strategies. These could reduce the significant mortality and morbidity of inflammation/infection, oxidative stress, insulin resistance, and immune modulation in metabolic disease associated with sPTL and PTB [122].

The use of machine learning using different statistical techniques such as model selection, Bayesian networks [123], mediation testing and imputation [124,125], which can correct for confounding variables, may compensate for traditional difficulties encountered with the complex interplay and comparisons between observational and interventional studies, demographic risk, and varying standards of care. These challenges and opportunities apply to many of the biomarkers discussed in this review, including transcriptomic, proteomic, immunological or microbiome markers.

1.12. Biobanking for preterm birth research

Biobanks are organized collections of mostly human biological material such as blood, urine, DNA and associated molecular, phenotypic, clinical, demographic and personal information and images stored for research purposes. Such biobanks have evolved from individual collections of biological specimens for investigator-led research to industry-supported biorepositories. Currently, biobanks are transforming the research landscape due to the availability of diverse, well-characterised biological specimens, often accompanied by a minimal dataset of anonymised demographic and clinical profiles, through to rich metadata as part of the drive towards improved biomarker discoveries and precision medicine.

Biobanks present many challenges, the most important of which is the sample integrity and analytical reproducibility so that procurement, processing, and preservation of specimens must be consistent. This is not an insignificant undertaking as biobank expansion continues especially when personnel, infrastructure and financial support is limited [130]. In addition to the upkeep, and operational and sustainability challenges, the ethics of biobanking remain under discussion. Individuals donating their samples do so with broad, informed consent, often without the need to specify the potential future use of the samples [131,132]. However, with accelerated scientific, technological advances, much of which may be unforeseen and/or unpredictable, the validity and ethics of retaining samples for every research eventuality comes into question. The ethical, legal, social implications flagged by biobanks in relation to sample storage usage and personal data has resulted in persuasive arguments being advanced for dynamic consent [132].

For pregnancy and perinatal research, dedicated biobanks remain relatively few. Those in existence are principally based in high-income countries with few in low-income or middle-income countries where the burden of PTB is highest [133]. Table 1 highlights some of the oldest established (CoLab, GAPPs) and most recent pregnancy biobanks (AMANHI, SAMBA, the Odense Birth Cohort and PRECISE [https://precisenetwork.org/]). In addition, there are several biobanks aligned
Development of novel biological markers based on extracellular vesicles, cell and cell free molecular component, along with establishment of biobanks around the globe, with standardized approaches, have substantially improved biomarker research and development in the past decade. Multiple governmental and other non-governmental organisations have contributed substantially to the establishment of biobanks in countries where PTB is prevalent, with annual rates >10%. The availability of standardized biological specimens has provided the opportunity to test and compare biomarkers in multi-ethnic cohorts. This has also helped to garner support from the Biotechnology industry and provided many new avenues for biomarker developmental projects.

Prior reports on PTB biomarkers such as pro-inflammatory cytokines, C-reactive protein and corticotrophin releasing hormone were primarily developed based on their functional relevance [40]. This led to marker selection bias and did not generate reliable biomarkers. The past decade has shown how unbiased approaches to assay and analysis, using well characterized biological specimens, primarily minimally invasive samples, can be used to discover combinations of biomarkers and validate them in multiple cohorts. In this report, we provide comment on recent developments using novel approaches in PTB biomarker research, emphasising their advantages and disadvantages, and future potential.

Biomarker discovery research and development often faces a valid question from clinicians such as, “do we need biomarkers when there are no interventions?” Lack of successful clinical trials to reduce the risk of sPTL can be attributed to factors such as: i) lack of understanding of the mechanisms by which a given drug may function; ii) improper selection of subjects who may benefit from a specific intervention due to lack of biomarkers to identify high risk status and iii) during which gestational window the intervention is best applied.

These factors suggest that successful interventions require identification of high-risk subjects using biomarkers indicative of underlying risk. Biomarkers may or may not indicate function and often bioinformatics analysis of a set of predictive markers can provide misleading information with respect to the disease process or underlying biological mechanisms and pathways. It is unlikely that universal biomarkers that will predict all high-risk pregnancies for sPTL will be developed and many may identify only specific subsets of cases. Nevertheless, we would like to argue that biomarkers, irrespective of their function, are products of an underlying physiological or pathophysiological process. Hence, biomarkers can generate hypotheses to test the causality in the generation of such a marker or markers in a given sample at a given time point from specific tissues during pregnancy. These studies have the potential to identify pathways and mechanisms and can lead to discovery of targets for intervention. Accordingly, we argue that successful biomarkers of PTB prediction are pathways to generate intervention and management strategies. This knowledge and specific biomarker profiles can also help to develop targeted or personalized interventions to reduce the risk of PTB.

3. Summary
In this Commentary, we have emphasised the likelihood that, in a clinical setting, the use of a combination of biophysical, biochemical, immunological, microbiological, fetal cell, exosomal, or cell free RNA at different gestational ages, may have to be implemented as part of a multivariable predictor model. Added to these will be the systems biological data using “omics” data. We appreciate that the data from these different scientific approaches using biobanked specimens together with data-banked information will require the use of bioinformatics including complex computational and statistical data handling, including methods from machine learning and artificial intelligence, to derive translatable messages from the complex “Big Data” assembled. Future biomarker research will necessitate a combination of clinical and basic science research, and attempts must be made to develop and introduce non-invasive as opposed to invasive tests. We must also move away from using PTB as a surrogate marker for predictive or preventive measures to studies of birth outcomes that store and curate samples across generations from mother to baby. Pregnancy or perinatal biobanks typically store maternal plasma, serum and urine, placental tissue and cord blood, but some also process and store paternal plasma [134]. Given the increased interest in the human microbiome, larger biobanks also house rectal and/or vaginal swabs. Many of the biobanks store tissues from a range of pregnancy disorders including PTB, preeclampsia, and intra-uterine growth restriction (IUGR).

For pregnancy focused research involving biobanks, PREBIC have published guidelines designed to promote consistency for carrying out “omics” research and developed standard operating procedures [135]. Research into the aetiology, prediction and prevention of PTB may significantly improve feto-maternal outcomes through properly constituted, shared biobanks, big-data and other resources.

2. Discussion
Preterm birth is the major cause of death and handicap in newborn babies worldwide. While the rate of sPTL and PTB globally has not decreased significantly over recent decades [136], there have been significant advancements in clinical care and biomedical research and development [93].

Over the past decade, the PREBIC organisation has been significantly influential worldwide in raising the profile of sPTL leading to PTB, particularly in the area of PTB biomarker research. Through a series of systematic reviews [40,137–139], PREBIC has exposed the knowledge gaps in biomarkers and various issues associated with PTB biomarker research [79]. Earlier reports highlighted many actions that were needed prior to the conduction of biomarker research and development. Those include, but are not limited to, homogeneity in the definition of the phenotype, geographic location of the cohort and study design, type of sample and sampling time such as gestational age, collection, processing, storage, assay technique, analysis and interpretation. The consensus is that, with a complex syndrome like sPTL leading to PTB, (as opposed to elective PTB for maternofetal indications), cannot be predicted by a single biomarker. Accordingly, the identification of multiple, or a combination of biomarkers at different gestational ages, will be required.

However, multi-“omics” and high-throughput technologies, the

<table>
<thead>
<tr>
<th>Sample of biobanks established for pregnancy research*</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMANHI <a href="http://www.everywomentogether.org/amandhimanoe">http://www.everywomentogether.org/amandhimanoe</a></td>
<td>Alliance for Maternal and Newborn Health Impr</td>
</tr>
<tr>
<td>Baby Biobank <a href="https://www.ucl.ac.uk/ch">https://www.ucl.ac.uk/ch</a></td>
<td>University College London and Imperial College London Collaborative to study four major pregnancy complications</td>
</tr>
<tr>
<td>Global Pregnancy Collaboration CoLab</td>
<td>International consortium of investigators and centers</td>
</tr>
<tr>
<td>GAPPs <a href="https://www.gapps.org/">https://www.gapps.org/</a></td>
<td>Global Alliance to prevent prematurity and stillbirth</td>
</tr>
<tr>
<td>MOMI <a href="https://mageewomens.org/">https://mageewomens.org/</a></td>
<td>Magee Obstetric Maternal and Infant database and biobank</td>
</tr>
<tr>
<td>SAMBA <a href="https://www.medscinet">https://www.medscinet</a> .com/samba/about.aspx?lang</td>
<td>Preterm Screening and Metabolomics in Brazil and Auckland</td>
</tr>
</tbody>
</table>

* The authors do not claim that this a complete list of biobanks.
Table 2

<table>
<thead>
<tr>
<th>Biomarker type</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA</td>
<td>Cheap and robust qPCR read-out</td>
<td>Time to result ranges from hours (qPCR assay) to days (sequencing assay)</td>
</tr>
<tr>
<td></td>
<td>Non-invasive</td>
<td>Discovery platform is very complex</td>
</tr>
<tr>
<td></td>
<td>Very sensitive, broad dynamic range</td>
<td></td>
</tr>
<tr>
<td>Proteomics</td>
<td>Proteins are close to phenotype and a measure of both existing and future health</td>
<td>Proteomic technology is complex, expensive and throughput limited</td>
</tr>
<tr>
<td></td>
<td>Protein measurement is a mainstay of clinical diagnostics</td>
<td>The proteome is vast and not yet approachable in a comprehensive way</td>
</tr>
<tr>
<td>Fetal cell</td>
<td>Both, quantification of fetal cells and expression of certain genes within the fetal cells can be used as biomarkers to predict PTB</td>
<td>Rarity of fetal cells in maternal blood circulation.</td>
</tr>
<tr>
<td></td>
<td>Potential to develop an automatic technology with low turn-around time to isolate fetal cells.</td>
<td>Number of fetal cells can be affected by other underlying conditions</td>
</tr>
<tr>
<td>Extracellular vesicles</td>
<td>Concentrate biomarker signal</td>
<td>Isolation techniques not standard</td>
</tr>
<tr>
<td></td>
<td>Tissue signal specificity</td>
<td>Concentrations can be low and near limit of detection</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>Currently, cervical length measurement is the only method that is commonly used in the clinical setting to identify women at risk for sPTB.</td>
<td>Cost-effectiveness analyses models have not consistently shown that universal cervical length screening is cost-effective</td>
</tr>
<tr>
<td></td>
<td>Only about 30% of women with a short cervix (who do not undergo an intervention) eventually deliver before 35 weeks gestation.</td>
<td></td>
</tr>
<tr>
<td>Microbiome</td>
<td>Molecular based tests have the advantage of:</td>
<td>Current evidence for the use of vaginal microbiome studies to predict PTB is sparse, less than robust, heterogeneous, conflicting, and of poor quality, such that a minimum database of recommendations for future studies has been proposed by PREBIC [79].</td>
</tr>
<tr>
<td></td>
<td>objectivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>quantification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>detection of fastidious organisms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>validity for self-obtained vaginal swabs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cost</td>
<td></td>
</tr>
<tr>
<td></td>
<td>speed of the test</td>
<td></td>
</tr>
<tr>
<td></td>
<td>accuracy</td>
<td></td>
</tr>
<tr>
<td>Fibronectin</td>
<td>Established bedside test</td>
<td>Not licensed for use in USA</td>
</tr>
<tr>
<td></td>
<td>Used in a validated algorithm in cervical length that assists clinical management of women at risk of preterm birth and those who presenting in threatened preterm labor</td>
<td>Cost of test limits widespread adoption particularly in LMIC</td>
</tr>
<tr>
<td></td>
<td>Has potential to be combined with other new biomarkers</td>
<td>Use restricted to &gt;18 weeks pregnancy</td>
</tr>
</tbody>
</table>

*Key points are listed; however, this list is not a comprehensive list.

more emphasis on the short and long term neonatal outcome of PTB according to different phenotypes of PTB. A summary of potential new biomarkers, their advantages and disadvantages at a glance, is provided Table 2. We would encourage cost-effectiveness analyses for future biomarker studies. Finally, we would like to emphasise the global nature of the problem of sPTB, that leads to PTB and encourage international collaboration in the field with shared databases, standard operating procedures and the use of minimum datasets.

Acknowledgements

We would like to thank Hologic, Inc. USA, who sponsored the PREBIC Biomarker summit, and the University of Southern Denmark for hosting the meeting in Odense, Denmark, (Nov 23–25, 2019).

References


