Paper:
Seale, A., Koech, A., Sheppard, A., Barsosio, H., Langat, J., Anyango, E., Mwakio, S., Mwarumba, S., Morpeth, S., Anampiu, K., Vaughan, A., Giess, A., Mogeni, P., Walusuna, L., Mwangudzah, H., Mwanzui, D., Salim, M., Kemp, B., Jones, C., Mturi, N., Tsofa, B., Mumbo, E., Mulewa, D., Bandika, V., Soita, M., Owiti, M., Onzere, N., Walker, A., Schrag, S., Kennedy, S., Fegan, G., Crook, D. & Berkley, J. (2016). Maternal colonization with Streptococcus agalactiae and associated stillbirth and neonatal disease in coastal Kenya. *Nature Microbiology, 1*(7), 16067

http://dx.doi.org/10.1038/nmicrobiol.2016.67

Maternal colonisation with *Streptococcus agalactiae,* and associated stillbirth and neonatal disease in coastal Kenya

Anna C Seale (DPhil),[1,2*] Angela C Koech (MBChB),[2] Anna E Sheppard (PhD),[3] Hellen C Barsosio (MBChB),[2] Joyce Langat (HND),[2] Emily Anyango (HND),[2] Stella Mwakio (HND),[2] Salim Mwarumba (MSc),[2] Susan C Morpeth (PhD),[2,4] Kirimi Anampiu (HND),[2] Alison Vaughan (BSc),[3] Adam Giess (MSc),[3] Polycarp Mogeni (MSc),[2] Leahbell Walusuna (MSc),[2] Hope Mwangudzah (HND),[2] Doris Mwanzui (RN),[5] Mariam Salim (RN),[2] Bryn Kemp (DPhil),[2,6] Caroline Jones (PhD),[1,2,4] Neema Mturi (MRCPCH),[2] Benjamin Tsofa (DPH),[2] Edward Mumbo (BScN),[7] David Mulewa (MBChB),[7] Victor Bandika (MMed),[8] Musimbi Soita (MMed),[9] Maureen Owiti (MMed),[6] Norris Onzere (MMed),[5] A Sarah Walker (PhD),[3] Stephanie J Schrag (DPhil),[10] Stephen H Kennedy (MD),[6] Greg Fegan (PhD),[1,2] Derrick Crook W (FRCPath),[3] James A Berkley (FRCPCH).[1,2]

1. Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, UK.
2. KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya.
3. Modernizing Medical Microbiology Consortium, Nuffield Department of Clinical Medicine, University of Oxford, UK.
4. London School of Hygiene and Tropical Medicine, London, UK
5. Maternity Department, Kilifi County Hospital, Kilifi, Kenya
6. Nuffield Department of Obstetrics and Gynaecology, University of Oxford, UK.
7. County Ministry of Health, Kilifi, Kenya.
8. Department of Paediatrics, Coast Provincial General Hospital, Mombasa, Kenya.
9. Department of Obstetrics and Gynaecology, Coast Provincial General Hospital, Mombasa, Kenya.
10. Division of Bacterial Diseases, Centers for Disease Control and Prevention, Atlanta, USA.

Abstract 198/200

Main Text 3496/3500

Abstract

Streptococcus agalactiae (Group B Streptococcus, GBS) causes neonatal disease and stillbirth, but its burden in sub-Saharan Africa is uncertain. We assessed maternal recto-vaginal GBS colonisation (7967 women), stillbirth and neonatal disease. Whole genome sequencing was used to determine serotypes, sequence types (ST), and phylogeny. We found low maternal GBS colonisation prevalence (934/7967, 12%), but comparatively high incidence of GBS-associated stillbirth and early onset neonatal disease (EOD) in hospital (0.91(0.25-2.3)/1000 births; 0.76(0.25-1.77)/1000 live-births respectively). However, using a population denominator, EOD incidence was considerably reduced (0.13(0.07-0.21)/1000 live-births). Treated cases of EOD had very high case fatality (17/36, 47%), especially within 24 hours of birth, making under-ascertainment of community-born cases highly likely, both here and in similar facility-based studies. Maternal GBS colonisation was less common in women with low socio-economic status, HIV infection and undernutrition, but when GBS-colonised, they were more likely colonised by the most virulent clone, CC17. CC17 accounted for 267/915(29%) of maternal colonising (265/267(99%) serotype III, 2/267(0.7%) serotype IV), and 51/73(70%) of neonatal disease cases (all serotype III). Trivalent (Ia/II/III) and pentavalent (Ia/Ib/II/III/V) vaccines would cover 71/73(97%) and 72/73(99%) of disease-causing serotypes respectively. Serotype IV should be considered for inclusion, with evidence of capsular switching in CC17 strains.

Introduction

A half of all child deaths (<5 years) worldwide are in Sub-Saharan Africa (sSA),[1] and a third of these deaths are in the neonatal period, from infection, preterm birth and neonatal encephalopathy.[1] Stillbirths likely equal neonatal deaths in number, and infections are a major contributor.[2] *Streptococcus agalactiae* (Group B Streptococcus, GBS) causes neonatal early and late onset disease (EOD, LOD), stillbirth,[3] and possibly contributes to preterm birth[4] and neonatal encephalopathy,[5] from ascending maternal genito-urinary colonisation (Supplementary Table 1 gives definitions). Whilst GBS emerged as the leading cause of EOD in the United States in the 1960s[6] and subsequently in Europe, in sSA, there remain major questions as to whether GBS commonly colonises pregnant women, causes stillbirth, or is an important cause of neonatal disease. Establishing this is essential, to inform potential preventive interventions. In resource-rich countries, reductions in EOD have followed the introduction of maternal microbiological or risk factor screening with intra-partum antibiotic prophylaxis (IAP).[7] However, there is uncertainty as to the feasibility of this approach in resource-poor settings, and there is no evidence of effectiveness of IAP in preventing GBS-associated stillbirth, or LOD. Antisepsis at delivery has been shown to be ineffective.[8] However, maternal vaccination may provide a feasible strategy to reduce GBS disease in resource-poor countries. A trivalent conjugate vaccine (serotypes Ia/Ib/III) has completed phase 2 clinical trials,[9] and a pentavalent vaccine is in development.[10]

Understanding which women are most likely to be GBS colonised could provide insight into both the emergence of GBS, and variation in reported prevalence of maternal GBS colonisation: Europe/United States 5-40%,[11,12] Africa 9-47% (Supplementary Table 2). Reported maternal risk factors for colonisation are conflicting, with increased maternal GBS colonisation reported in both younger[13] and older[14] age groups; African-American mothers,[13-15] and those with higher

education,[14,16] higher income,[16] high sexual activity,[14] and obesity.[15,16] Data from sSA are limited, but are also conflicting for potentially important risk factors such as HIV-infection. In South Africa, maternal GBS colonisation was lower in HIV-infected mothers[17] but in Malawi, only amongst HIV-infected mothers with lower CD4 counts.[18] In the USA[15] and Zimbabwe[19] no association with HIV was found. The limited data from studies in Kenya, Zimbabwe, Malawi and South Africa on colonising maternal serotypes in sub-Saharan Africa suggest serotype III is the most common (Ia/Ib/II/IV/V also reported).[18,20-22]

For neonatal disease, data outside of the United States and Europe are sparse.[23] In sSA, facility-based studies generally report a high incidence of neonatal GBS disease, but population-based and outpatient studies have reported much lower incidences,[24,25] including what was described as a “striking absence” of invasive neonatal GBS disease in large out-patient based studies.[24] However, regional estimates, that included only four studies from Africa (one of which is our study site in Kilifi County)[8,26-28] suggest that Africa may have the highest regional burden of neonatal GBS disease at 1.2(0.50-1.91)/1000 live-births.[23] These limited data suggest that serotype III, as described in other regions,[23] most commonly causes disease; for EOD and LOD in Malawi 52% and 72%;[27] in South Africa 49% and 76%,[29] with serotypes Ia/Ib/II/V also reported.[27,29] The incidence of GBS-associated stillbirth is unknown in sub-Saharan Africa,[3] with data from two studies; one found no GBS-associated stillbirth,[30] the other 8/66(12%) stillbirths.[31]

The population structure of GBS in Europe and the United States can be described by five major clonal complexes: CC1, CC10, CC17, CC19 and CC23,[32,33] with CC17 overrepresented in disease isolates.[32,34] These five clonal complexes are also found in Africa;[32] in addition, CC26 is common in some regions, representing 15% of sampled GBS isolates in Dakar and Bangui.[35]

GBS also causes bovine mastitis, which is largely mediated by the bovine-specific CC67, although the five major human clonal complexes can also be found in cattle.[33,36,37]

In this study, we aimed to comprehensively describe the clinical epidemiology of maternal GBS colonisation, neonatal disease and stillbirth in coastal Kenya, with molecular analysis to determine associated serotypes, sequence types (ST), and phylogeny.

Results

Maternal GBS colonisation and adverse perinatal outcomes

During the study, 10,130 pregnant women attended a health facility and we recruited 7,967 (Figure 1, sample size Supplementary Table 3). Of these, 526/7967(6.6%) were from rural sites, 5470/7967(68.7%) from semi-rural and 1971/7967(24.7%) from an urban site. There were some differences in demographics in those excluded (Supplementary Table 4), with emergency referrals more likely to be excluded as well as women with incomplete data on age, ethnicity or parity, although overall numbers were small. Transport times to the laboratory were longer from urban and rural sites (median 11h(range 0-48h); 11h(0-52h) respectively compared to semi-rural (5h(0-73h)), but there was no evidence of association between GBS isolation and time to sample processing across all sites (OR=1.00(0.99-1.00) p=0.6)), across rural and urban sites (OR=0.99(0.98-1.00)), or each site individually (Supplementary Figure 1).

Overall, 934 (11.7%(11.0-12.5%)) women were GBS-colonised at delivery. Prevalence was lowest at the rural sites (47/526, 8.9%(6.6-11.7%)), intermediate in the semi-rural site (608/5470, 11.1%(10.2-12.0%)) and highest at the urban site (279/1971, 14.2%(12.6-15.8%); trend P<0.001). However, after adjustment for other risk factors (including maternal age, socio-economic status and ethnicity; univariable analyses Supplementary Table 5), the odds of

isolating GBS at the urban site (OR=0.95(0.92-0.98)) and rural site (OR=0.91(0.88-0.94)) were lower than at the semi-rural site (p<0.001), Table 1.

GBS colonisation was independently associated with maternal age, highest in the middle categories (Supplementary Figure 2; p=0.023), and parity (≥5 vs 1-4) (OR=0.81(0.70-0.93) p<0.001) as well as Mijikenda ethnicity (indigenous population, OR=0.73(0.59-0.90) p=0.003) (Table 1). GBS colonisation was increased in women with higher socio-economic status (OR=1.21(1.13-1.29), p<0.001) and those who had contact with cattle (OR=1.29(1.17-1.43) p<0.001). GBS colonisation was reduced amongst HIV-infected women, and especially in HIV-infected women taking co-trimoxazole prophylaxis (OR=0.68(0.42-1.09); OR=0.24(0.14-0.39), p<0.001), in less well-nourished mothers (OR=0.72(0.60-0.88), p<0.001) and women with obstetric emergencies (OR=0.85(0.79-0.92), p<0.001).

There was evidence that adverse perinatal outcomes (very preterm delivery, very low birth-weight, stillbirth, possible serious bacterial infection (definitions Supplementary Table 1) were associated with maternal GBS colonisation in multivariable models in the context of interactions with clinical risk factors for invasive GBS disease, such as maternal temperature >37.5°C, urinary tract infection, and prolonged rupture of membranes >18h (Figure 2, Supplementary Tables 6-9). In contrast, without GBS colonisation there was no evidence that these clinical factors conferred elevated risk of poor outcomes. There was no evidence of association of maternal GBS colonisation with perinatal mortality (p=0.7; Supplementary Table 10), including testing for an interaction with any risk factor for GBS disease (p=0.4).

Of 918/934(98.3%) colonising isolates available and extracted, 915/934(98.0%) were of sufficient quality for genomic analysis. Amongst colonised mothers, 658/915(71.9%) of GBS isolates were serotypes Ia, Ib or III; serotype III being most common (350/915(38.3%)); Clonal-complex 17 (CC17) comprised 267/915(29.2%), Figures 3 and 4, Supplementary Table 11,

GBS- colonised women. Of these, 265/267(99.3%) were serotype III and 2/267(0.7%) were serotype IV.

The population structure was broadly similar to other parts of the world, with 114/915(12.5%) CC1, 148/915(16.2%) CC10, 268/915(29.3%) CC17, 173/915(18.9%) CC19, 208/915(22.7%) CC23, whilst 4/915(0.4%) did not belong to any commonly described clonal complex. No bovine-associated CC-67[38] GBS isolates were identified. Each of the five major clonal complexes were represented at each site (Figure 4, Supplementary Table 12), with no evidence for geographic stratification. Within clonal complexes, there was considerable diversity, with a total of 43 distinct STs, 18 of which were newly identified in this study. The largest number of STs was seen in CC17 (12 STs total, 8 newly identified). The most common STs within CC17 were ST17 (183/268,68.3%) and ST484 (67/268,25.0%).

Within GBS-colonised women, risk factors for colonisation with the most virulent clone CC17, were, in general, the reverse of those associated with GBS colonisation overall (Table 2). Maternal GBS CC17 was increased in the rural site (OR=1.26(1.20-1.31), p<0.001), women of Mijikenda ethnicity (OR=1.62(1.43-1.85), p<0.001), and women with HIV-infection and women with HIV-infection taking co-trimoxazole (OR=1.46(1.11-1.92); OR=4.30(0.59-31.3), p<0.001). Mothers who had cattle contact (OR=0.54(0.45-0.64), p<0.001) and were better nourished (OR=0.79(0.42-1.49), p<0.001) were less frequently colonised with CC17, but this did not hold for ST-17 (Supplementary Table 13). For each of the risk factors, including cattle contact, the corresponding isolates were dispersed in the phylogeny (Figure 4), suggesting that the associations were not driven by specific sub-lineages.

Pairwise comparison of all maternal colonising isolates in mothers delivering at Kilifi County Hospital showed increased genetic similarity in a small number of mothers who delivered within 7 days of each other, but not according to household location (Supplementary Figure 3). Of

mothers admitted <7 days apart, in Kilifi County Hospital, there were 14/91013(1.4%) pairs from mothers admitted on the same day with 0-4 Single Nucleotide Variant (SNV) differences, 11/1967(0.6%) 1 day apart, 2/1845(0.1%) 2 days apart and 2/1832(0.1%) 6 days apart (p<0.001). At the rural sites, of mothers admitted <7 days apart, there were 2/124(1.6%) pairs from mothers admitted on the same day with 0-4 SNV differences and 2/219(0.9%) 1 day apart (p=0.1). At the urban site, there were 8/987(0.8%) pairs from mothers admitted on the same day with 0-4 SNV differences and 3/1555(0.2%) 1 day apart (p<0.001).[22]

GBS in mother-neonatal pairs (surface contamination)

We recruited 830 mother and baby pairs at KCH (Figure 1, and Supplementary Table 14); 104/830 (12.5%(10.4-15.0%)) mothers were colonised with GBS at delivery and 44/830 (5.3%(3.9-7.1%)) neonates had GBS isolated from ear, umbilicus or nose within 6h of delivery. 30/44(68.2%) neonates with surface GBS were born to one of the 104 GBS-colonised mothers and 14/44(31.8%) were born to one of the 726 mothers without colonising GBS detected; of which 2/14(14.3%) were born by caesarean section. Odds of neonatal surface GBS were high with maternal GBS colonisation (OR=20.6(10.5-40.6,p<0.001)).

Pairwise SNV comparisons between maternal and newborn isolates showed a clear bimodal distribution: 26/30(86.7%) pairs differed by ≤4 SNVs (all pairs the same ST and serotype), presumably representing vertical transmission, and 4/30(13.3%) pairs were highly divergent (>9000 SNVs, with different STs and different serotypes), Figure 4. Combining all pairs with ≤4 SNVs, the SNVs were dispersed throughout the genome, with no gene represented more than once. There were 7/44(15.9%) neonates with surface GBS after delivery by caesarean section, 5 of their mothers had GBS detected; 3/5 had 0 SNV differences, 1/5 1 SNV, and 1/5 9673 SNVs.

Stillbirth

There were 278 stillbirths during the nested case-control study (278/4394(6.3%) all births). We sampled cord blood in 149/278(53.6%) (94/149(63.1%) intra-partum, 55/149 (36.9%) ante-partum stillbirths) 104 also had a lung aspirate; 34/278 (12.2%) had a lung aspirate sample only. In total 183/278(65.8%) stillbirths were sampled, plus 330 live-birth cord blood controls (Figure 1).

GBS was isolated from 4/183 (2.2%(95%CI0.6-5.5)) stillbirths (3/149 cord blood samples, 2/138 lung aspirates; one stillbirth had GBS isolated from both); two ante-partum (36 and 39 weeks' gestation) and two intra-partum (35 and 39 weeks'). Overall minimum incidence of GBS-associated stillbirth (cord blood or lung aspirate) was 0.91(0.3-2.3)/1000 births. Compared to live-born controls (GBS isolated from 1/330(0.3%)), GBS was isolated more frequently from cord-blood in stillbirths (OR=6.8(0.7-65.5), p=0.09), and in a multinomial model ante-partum stillbirths (OR=12.4(1.1-139.3)) and intra-partum stillbirth (OR=3.5(0.2-57.1) exact p=0.055). Serotype data were available from three stillbirths; two were serotype V and one serotype III.

There were 2/4 GBS-associated stillbirths born to GBS colonised mothers (2/2 pairs differed by 0 SNVs, all ST1, serotype V); one mother was not colonised, one was not tested. Risk ratio for GBS-associated stillbirth in GBS-colonised vs non-colonised mothers 7.6(1.1-52.6, p=0.016).

Neonatal disease

Eighty-two neonates with invasive GBS disease were admitted to KCH (1998-2013, Figure 1): 36/82(43.9%) and 43/82(52.4%) with EOD and LOD respectively (3 unknown). Case fatality was highest in EOD 17/36(47.2%) despite treatment, particularly for those diagnosed <24h of birth (11/18(61.1%)). In cases of LOD, 5/43(11.6%) died. Most GBS EOD cases (52/82(63.4%)) were male, and 25/82(30.5%) were <2500g at admission (Supplementary Table 15). Sepsis without

focus was predominant in EOD (33/36(91.6%)), with meningitis (+/- sepsis) being more common in LOD (21/43(48.8%)), (Figure 3). Gestational age was not routinely available from prior clinical surveillance data, however, there were five EOD cases with gestations of 36, 36, 37, 37 and 40 weeks' born at the time of the prospective cohort study (vs median 38 (IQR 36-40) overall in prospective cohort).

EOD incidence amongst deliveries at KCH during the cohort study (2011-13) was 0.76(0.25-1.77)/1000 live-births. Including only residents in KHDSS population (1998-2013), the (minimum) population-based incidence of neonatal GBS disease was 0.34(0.24-0.46)/1000 live-births: EOD 0.13(0.07-0.21)/1000 live-births and LOD 0.21(0.14-0.31)/1000 live-births; with no evidence of a trend over the study period (Supplementary Figure 4).

There were 73/82(89.0%) neonates with invasive isolates available and extracted, and all were of sufficient quality for inclusion in the final analysis. Serotypes Ia/Ib/III caused 71/73(97.3%) and serotypes Ia/Ib/II/III, caused 72/73(98.6%) of EOD and LOD. Serotype III predominated in both EOD (18/30(60.0%)) and LOD (36/40(90.0%); p=0.003 χ^2test for trend); these isolates were all CC17, except 1 CC-19 isolate (Figure 4). Serotype III was the almost universal cause of meningitis; 22/23(95.7%) cases, of which 21/22(95.4%) were CC17; Figure 3, Supplementary Table 16. Isolates were all susceptible to penicillin and 61/76(80.3%) were susceptible to co-trimoxazole.

Three of the five neonates with EOD born at KCH (2011-2013) were born to GBS-colonised mothers (1/3 pairs differed by 0 SNVs (both ST17, serotype III), 1/3 88 SNVs (1 ST17, 1 ST484, both serotype III) and 1/3 1002 SNVs (both ST17, serotype III): risk ratio (RR) for EOD for GBS-colonised vs non-colonised mothers 11.8(2.0–70.3) p<0.001. For all perinatal GBS disease (EOD or stillbirth) RR=13.1(3.1–54.8, p<0.001).

Discussion

GBS is an important cause of stillbirth and neonatal disease in Kenya. The incidence of stillbirth was comparable to early onset disease (EOD) in hospital births ((0.91(0.25-2.3)/1000 births) and 0.76(0.25-1.77)/1000 live-births respectively). These incidences are all underestimates, with samples not taken from all stillbirths, and insensitivity in cultures, particularly if intrapartum antibiotics were given. The much lower population-based incidence of EOD (0.13(0.07-0.21)/1000 live-births) suggests recruitment bias with under ascertainment of cases in the community, or in out-patient settings, due to rapid case fatality after delivery and limited access to care. This is supported by the higher proportion of late onset disease (LOD), which is the reverse of the ratio of GBS disease typically seen in high-income countries.[23] Whilst it could be argued that facility delivery is a risk factor for EOD (if there was in-hospital maternal GBS acquisition), we found very limited evidence of horizontal transmission in facilities, with few genetically near-identical pairs (0-4 SNVs, threshold determined empirically from newborn surface contamination study) in mothers admitted <7days of each other.

However, there may be true differences in incidence of both GBS-associated stillbirth and neonatal GBS disease in sub-Saharan Africa, neither explained by study design nor other methodological limitations. The incidences of neonatal GBS disease recently reported in urban South Africa[29] and Malawi[23] are high, and could be due to differences in maternal GBS colonisation prevalence; consistent with our finding of higher prevalence of maternal GBS colonisation in urban compared to semi-rural and rural residents. This association was explained by variables describing improved socio-economic status, and other factors associated with improved health, such as better nutritional status, being in the middle age categories, and lower parity, both in the complete-case analyses and using multiple imputation. Whilst our study

includes impoverished populations, the pattern of risk factors identified is consistent with recent studies in high-income countries reporting increased maternal GBS colonisation with higher education[14,16] and higher income.[16] The reasons for this are unclear, but it likely relates to changes in the maternal microbiome, with different community-states reported.[39]

Use of prophylactic co-trimoxazole amongst HIV-infected women had a clear negative association with GBS colonisation. Previously reported conflicting findings,[17,18] may depend on the frequency of antimicrobial use (and provision of anti-retroviral therapy). In contrast, neonatal GBS disease is increased with HIV-exposure,[40] with reduced maternal GBS capsular antibody in HIV-1 infection,[41,42] and/or because, as shown here, the most virulent clone, CC17, is more frequently found in HIV-infected GBS colonised women, compared to other non-CC17 types. There have been a number of virulence factors (adhesins, invasins and immune evasins) associated with increased ability of GBS to colonise and cause disease,[43] with the more homogeneous CC17 having acquired its own set of virulence genes,[38] and increased ability to form biofilms in acidic conditions.[44]

We observed an association between cattle contact and maternal GBS colonisation, however, no bovine-associated CC-67 isolates were identified, and the isolates from women with cattle contact were from a variety of lineages representing all major CCs. Little is known about bovine GBS populations in Kenya, and it is possible that the human and bovine populations are similar, and thus the association between cattle contact and maternal GBS colonisation from genuine transmission, as suggested elsewhere.[45] Alternatively, women who look after cattle may be of higher socio-economic status and thus the association due to residual confounding.

The overall GBS population structure here is similar to previous studies from a variety of geographic locations, supporting the notion of recent global dissemination of relatively few clones.[32] Within this study, we found no evidence for geographic clustering of related isolates,

both at the level of sampling location (Figure 4), as well as distance between households (Supplementary Figure 3), further suggesting rapid geographic dispersal of GBS. However, in contrast to a previous study from Africa,[35] we found no CC-26 isolates, suggesting this lineage may be geographically restricted. Furthermore, we found a large number of ST-484 isolates 67/915(7.3%) of total, 67/268(25.0%) of CC17; this lineage has previously been reported in only a single study, also from Kenya.[46] We also identified three novel STs that represent single-locus variants of ST-484. Taken together, it is possible that ST-484 originated in or near Kenya, with relatively little geographic dispersal. Alternatively, there may be a lack of GBS sampling in other locations where ST-484 is present.

Prevention strategies in resource-rich settings focus on reducing EOD through intra-partum antibiotic prophylaxis (IAP) using either microbiological or risk-factor screening to identify at-risk mothers;[7] both strategies would be challenging in resource-poor settings. Of interest, when comparing these strategies, however, is the fact that associations with adverse perinatal outcomes were only detected through interactions between maternal GBS colonisation and clinical risk factors. This supports a mechanism of action whereby colonising maternal GBS ascends, leading to chorioamnionitis (intra-amniotic infection) and fever in a small proportion of women, leading to poor perinatal outcomes. Neither maternal GBS colonisation without signs of infection, nor maternal fever without GBS colonisation increased the risk of adverse perinatal outcomes. Thus either approach (microbiological or risk-factor screening) will target far larger numbers than those actually at risk. Any direct association between maternal GBS colonisation and adverse outcomes may also be diluted by the many other causes of adverse perinatal outcomes, and by misclassification (e.g. uncertainty over the date of the last menstrual period to determine gestation), which may explain some of the conflicts in findings in studies assessing the contribution of GBS to preterm birth.[4]

We demonstrated vertical transmission of maternal GBS colonisation in maternal-newborn dyads, for both surface contamination (including in cases of emergency caesarean section) and perinatal disease. Genetically divergent maternal-newborn dyads may reflect un-sampled variation in the mother, as only a single colony was sequenced in each case. Whilst adaptive mutations associated with disease progression have been reported elsewhere from the comparison of mother-newborn pairs,[47] we were unable to find evidence for this in the current study, as all pairs involving invasive isolates were either genetically identical (0 SNVs), or divergent enough to argue against this. The findings show GBS infection occurs prior to delivery; supporting the need for IAP to be administered before delivery to be effective, and showing why antisepsis in active labour, for example vaginal chlorhexidine wipes, are ineffective in reducing neonatal EOD.[8] The finding of 14/44(31.8%) newborns with surface GBS contamination, where maternal GBS colonisation was not identified suggests insensitivity of maternal recto-vaginal screening, despite the consistent use of broth-enrichment and blood agar to maximise sensitivity. This is a higher percentage than a recent study in The Gambia (40/186(21.5%)),[48] but this study excluded mothers at high risk for pregnancy complications. Similarly to repeat vaginal examinations, as seen here and reported elsewhere,[49] complicated deliveries (obstetric emergencies) likely decrease GBS sampling sensitivity, through antisepsis measures, or mechanical removal.

With limitations in the clinical benefit of IAP in terms of reducing stillbirth and LOD, as well as challenges in effective implementation to reduce EOD in sSA, maternal vaccination is an attractive strategy for prevention. The most advanced vaccine (completed phase 2 trials) is trivalent (Ia/Ib/III), but plans are to advance a pentavalent vaccine.[10] If this includes the most common disease-causing serotypes worldwide (Ia/Ib/II/III/V), it will cover almost all 72/73(98.7%) of the serotypes causing invasive disease in this study. However, importantly for vaccine development, and in line with other reports,[50] we identified capsular switching to

serotype IV in 2 isolates within CC17, suggesting consideration of inclusion of serotype IV is warranted.

GBS is an important, potentially preventable, cause of stillbirth and neonatal death in coastal Kenya. Maternal GBS colonisation is increased with urbanisation and higher socio-economic status, and likely to increase with development. GBS neonatal disease in population-based studies is markedly under-ascertained through rapid case fatality after birth and limited access to care, and is equalled by the burden of GBS-associated stillbirth. Maternal GBS vaccination is a key opportunity to reduce stillbirth and neonatal death in this high burden region.

Methods

Study design

The study design included a prospective cohort at rural, semi-rural and urban sites, a nested case-control study in the semi-rural site, and analysis of surveillance of neonatal disease at the semi-rural site (Figure 1).

Prospective cohort study: In a prospective cohort study (2011-13), we assessed prevalence and risk factors for maternal GBS colonisation at delivery, and perinatal outcomes at delivery (stillbirth, gestational age, birth-weight, possible serious bacterial infection, and perinatal death).

Nested case control study: Investigation of stillbirth was undertaken with a nested case-control study; Cord blood cultures were taken at delivery from the stillbirth, and the next two subsequent admissions that were live-born (case: controls 1:2). Lung aspirates were taken from stillbirths only, by a study clinician attending within 4 hours of the stillbirth.

Surveillance of neonatal invasive bacterial disease: Neonatal disease was quantified using systematic clinical and microbiological surveillance data (1998-2013 at Kilifi County Hospital)

within the Kilifi Health and Demographic Surveillance System (KHDSS) area, giving accurate population and birth denominators (see study sites).[51]

Study sites

The studies were conducted at Coast Provincial General Hospital, Mombasa (CPGH) (urban, ~12,000 deliveries/year, comprehensive obstetric care); Kilifi County Hospital (KCH) (semi-rural, ~3000 deliveries/year, comprehensive obstetric care); Bamba sub-district hospital (rural, ~600 deliveries a year, basic obstetric care) and Ganze health facility (rural, ~400 deliveries a year, basic obstetric care).

A part of Kilifi County is included in detailed health and demographic surveillance (KHDSS)[51] from which accurate population data are available from 2004. Kilifi County Hospital (KCH) is the main district hospital which serves this population, so incidence estimates for residents seeking health care at KCH can be made with the KHDSS population as the denominator. We used prospectively collected data on live births from the regular re-enumerations of the KHDSS population, and used the estimated slope from a regression to estimate the number of births prior to the start of KHDSS.

Study population

Prospective cohort study: We included all women admitted for delivery at study sites admitted at designated times who gave written informed consent, without additional exclusion criteria. We planned to recruit over one calendar year (to allow for seasonality), but extended enrolment to meet sample size requirements (Supplementary Table 3) because national strikes closed government health facilities twice during the study. Recruitment was done at CPGH for 48 hours each week (01.04.2012-31.07.2013), at Bamba and Ganze for 6 days each week (01.07.2012-31.07.2013) and at KCH every day (01.08.2011-31.07.2013) including additional studies of neonatal surface contamination (01.05.2012 to 31.07.2013)

Nested case control study: We included all stillbirths delivered in Kilifi County Hospital and the next two consecutive live births (01.05.2012-01.10.2013).

Surveillance of neonatal invasive bacterial disease: We included all neonates admitted to Kilifi County Hospital (01.08.1998-1.10.2013).

Sampling and laboratory methods

Prospective cohort study: We took recto-vaginal swabs during routine vaginal examination at admission for delivery, when possible prior to rupture of membranes. A small cotton swab was used to wipe the lower third of the vaginal mucosa and then the inside surface mucosa of the anus,[52] according to standard procedures. Neonatal surface swabs (to assess surface contamination) included the external ear, nares and umbilicus. Swabs were placed into Amies transport medium with charcoal,[53] refrigerated, transported in cool containers[53] to the research laboratory (participating in UK National External Quality Assessment Service) and processed by standard protocols (including enrichment (LIM broth) and sub-culture onto blood agar). Isolates with GBS morphology were CAMP tested and definitive grouping done using a Streptococcal grouping latex agglutination kit (PRO-LAB Diagnostics, USA).

Nested case control study: For stillbirths and live-born controls, we sampled cord blood at delivery after double clamping the cord if necessary and cleaning with 70% ethanol. We processed cord blood cultures using an automated culture system (BACTEC 9050, Becton Dickinson, UK). We took lung aspirate samples (stillbirths only) with a sterile technique aspirating the lung, within four hours of delivery. We examined lung aspirates with microscopy and culture using standard methods within 30 minutes of sampling, or if delay was unavoidable stored at 2-8^{o}C for up to 8 hours.

Surveillance of neonatal invasive bacterial disease: For all neonatal admissions (1998-2013) at KCH, we sampled peripheral blood on admission for culture, prior to neonatal antibiotic

403 treatment (during 2011-2013, peri-partum maternal antibiotics were documented in
404 36/5430(0.7%) of deliveries in KCH); we did lumbar puncture when clinically indicated. We
405 tested isolates for antimicrobial susceptibility to penicillin and co-trimoxazole (British Society for
406 Antimicrobial Chemotherapy). We processed blood cultures using an automated culture system
407 (BACTEC 9050); we tested cerebrospinal fluid as described elsewhere.[26]

408 **Molecular methods**

409 We performed DNA extraction, Illumina sequencing (Hiseq technology) and raw read
410 processing using standard methods starting from a single GBS colony. GBS isolates were
411 frozen in 1mL vials and stored at -80°C prior to sub-culture on a Columbia blood agar plate for
412 24-48 hours, followed by DNA extraction using a commercial kit (QuickGene, Fujifilm, Tokyo,
413 Japan) from a single colony. High throughput sequencing was undertaken at the Wellcome
414 Trust Centre for Human Genetics (Oxford University, UK) using HiSeq2500, generating 150
415 base paired-end reads. *De novo* assembly, mapping and variant calling were performed as
416 previously described,[54] except that mapping was to the *S. agalactiae* reference genome
417 2603V/R (NC_004116.1). Sequence quality was assessed using various metrics (% reads
418 mapped to reference genome, % reference positions called, contig number, total contig length).
419 Sequence data showing poor quality metrics was excluded from further analysis; where
420 practicable the corresponding samples were re-isolated, re-grouped and re-sequenced (if re-
421 grouping confirmed the isolate as GBS). Sequence data were submitted to the NCBI
422 Sequence Read Archive under BioProject PRJNA315969. Individual accession numbers are
423 provided in Supplementary Table 17 (BioProject PRJNA315969).

424 We allocated serotype on the basis of BLASTn comparisons assessing sequence similarity of
425 *de novo* assemblies with the capsular locus regions of each of the ten known GBS serotypes.
426 We validated this method internally (kappa=0.92).[55] Sequence types (ST) were also assigned *in*

silico using BLASTn with *de novo* assemblies. Novel STs were submitted to pubmlst.org for assignment. Phylogenetic analysis was performed separately for each clonal complex using RAxML version 8.1.16, with an alignment consisting of all variable sites from mapping to the 2603V/R reference, padded to the length of the reference with invariant sites of the same GC content as the original data. Recombination was detected using ClonalFrameML,[56] and we present the resultant phylogenies with recombinant regions removed. To partition the isolates according to previously described clonal complexes, we first reconstructed a single RAxML phylogeny with all isolates. The resulting tree was then visually partitioned on long, deep branches, which effectively corresponded to previously described clonal complexes, but enabled us to include all STs. We have therefore used this partitioning as our definition of the clonal complexes. Using this definition, each ST belongs to a single clonal complex and each clonal complex is monophyletic (Supplementary Figure 6), indicating that partitioning by clonal complex remains appropriate when whole-genome data is taken into account.

Pairwise comparison of SNV differences from mapped data was used to examine maternal and newborn paired GBS isolates, and possible transmission of GBS between mothers was investigated through these differences and epidemiological links in time and place (through delivery in Kilifi County Hospital) or residence (distance between household locations in Kilifi HDSS).

Statistical analysis

We used Stata (version 13.1) for statistical analyses. We used the first principal component from a set of household assets as a proxy for socio-economic status (SES).[57] We used multiple imputation with chained equations (Stata mi) to impute missing data on potential risk factors (<15% per variable; 50 imputations). Continuous variables were checked for normality and transformation was not required. We used natural cubic splines to allow for non-linearity in

variable effects in imputation models. Imputations were done separately by maternal GBS status so that interactions could be examined in the analyses of adverse newborn outcomes. The same imputation was used for both analyses; by imputing separately for GBS colonisation there are fewer assumptions than if it was fitted as a covariate (allows variances of continuous imputed variables to differ according to GBS colonisation, and the associations between two imputed variables can be stronger in one group).

We built multivariable logistic regression models using complete-case and imputed datasets (combined using Rubin's rules) to examine risk factors for maternal GBS colonisation using robust variances reflecting clustering by site. We included non-linearity in continuous variables via natural cubic splines, with factors categorised at quartiles for presentation of final models. Risk factors with $p<0.1$ in univariable models were included in a multivariable model and final independent predictors identified using backwards elimination (exit $p>0.1$). We assessed whether risk factors for maternal GBS colonisation were associated with ST-17 (and CC17) colonisation in mothers who were GBS colonised using the same process, for complete-cases only.

We used the imputed dataset in multivariable regression analyses to examine whether maternal GBS colonisation was associated with gestational length, birth-weight, possible serious bacterial infection, stillbirth or perinatal mortality. We included pre-specified confounders (age, parity, sex (of new-born), maternal education, SES, nutritional status, HIV status, obstetric complication and multiple delivery) and tested for interaction with GBS colonisation from prolonged rupture of membranes (PROM, >18h), maternal fever (>37.5°C) or urinary tract infection (leukocytes and nitrites present). We included these terms in multivariable models if there was evidence of interaction at the $p<0.1$ level.

We estimated the odds of isolating GBS from cord blood in all stillbirths, then ante-partum and intra-partum stillbirths, compared to live-births. We estimated incidence of GBS-associated stillbirth and neonatal disease using denominators of facility births, and community births, for residents of Kilifi Health and Demographic Surveillance Study.[51]

Ethics

The study protocol was approved by KEMRI Ethical Review Committee (SSC/ERC 2030) and the Oxford Tropical Research Ethics Committee (53-11) (clinicaltrials.gov NCT01757041).

References

1 Liu, L. *et al.* Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. *Lancet* **385**, 430-440, doi:10.1016/S0140-6736(14)61698-6 (2014).

2 Goldenberg, R. L., McClure, E. M., Saleem, S. & Reddy, U. M. Infection-related stillbirths. *Lancet* **375**, 1482-1490, doi:10.1016/S0140-6736(09)61712-8 (2010).

3 Nan, C. *et al.* Maternal group B Streptococcus-related stillbirth: a systematic review. *BJOG : an international journal of obstetrics and gynaecology* **122**, 1437-1445, doi:10.1111/1471-0528.13527 (2015).

4 Valkenburg-van den Berg, A. W., Sprij, A. J., Dekker, F. W., Dorr, P. J. & Kanhai, H. H. Association between colonization with Group B Streptococcus and preterm delivery: a systematic review. *Acta Obstet Gynecol Scand* **88**, 958-967, doi:10.1080/00016340903176800 (2009).

5 Tann, C. J. *et al.* Prevalence of bloodstream pathogens is higher in neonatal encephalopathy cases vs. controls using a novel panel of real-time PCR assays. *PLoS One* **9**, e97259, doi:10.1371/journal.pone.0097259 (2014).

6 Baker, C. J., Barrett, F. F., Gordon, R. C. & Yow, M. D. Suppurative meningitis due to streptococci of Lancefield group B: a study of 33 infants. *J Pediatr* **82**, 724-729 (1973).

7 Schrag, S., Zywicki S, Farley MM, Reingold AL, Harrison LH, Lefkowitz LB, Hadler JL, Danila R, Cieslak PR, Schuchat A. Group B streptococcal disease in the era of intrapartum antibiotic prophylaxis. *N Engl J Med* **342**, 15-20 (2000).

8 Cutland, C., Madhi SA, Zell ER, Kuwanda L, Laque M, Groome M, Gorwitz R, Thigpen MC, Patel R, Velaphi SC, Adrian P, Klugman K, Schuchat A, Schrag SJ; PoPS Trial Team. Chlorhexidine maternal-vaginal and neonate body wipes in sepsis and vertical transmission of pathogenic bacteria in South Africa: a randomised, controlled trial. *Lancet* **374**, 1909-1916 (2009).

9 Heyderman, R. S. *et al.* Group B streptococcus vaccination in pregnant women with or without HIV in Africa: a non-randomised phase 2, open-label, multicentre trial. *Lancet Infect Dis*, doi:10.1016/S1473-3099(15)00484-3 (2016).

10 Mullard, A. Making way for maternal immunization. *Nature reviews. Drug discovery* **15**, 3-4, doi:10.1038/nrd.2015.38 (2015).

11 Barcaite, E. *et al.* Prevalence of maternal group B streptococcal colonisation in European countries. *Acta Obstet Gynecol Scand* **87**, 260-271, doi:10.1080/00016340801908759 (2008).

12 Schuchat, A. & Wenger, J. D. Epidemiology of group B streptococcal disease. Risk factors, prevention strategies, and vaccine development. *Epidemiologic reviews* **16**, 374-402 (1994).

13 Anthony, B. F., Okada, D. M. & Hobel, C. J. Epidemiology of group B Streptococcus: longitudinal observations during pregnancy. *J Infect Dis* **137**, 524-530 (1978).

14 Regan, J. A., Klebanoff, M. A. & Nugent, R. P. The epidemiology of group B streptococcal colonization in pregnancy. Vaginal Infections and Prematurity Study Group. *Obstet Gynecol* **77**, 604-610 (1991).

15 Shah, M., Aziz, N., Leva, N. & Cohan, D. Group B Streptococcus colonization by HIV status in pregnant women: prevalence and risk factors. *Journal of women's health* **20**, 1737-1741, doi:10.1089/jwh.2011.2888 (2011).

16 Stapleton, R. D., Kahn, J. M., Evans, L. E., Critchlow, C. W. & Gardella, C. M. Risk factors for group B streptococcal genitourinary tract colonization in pregnant women. *Obstet Gynecol* **106**, 1246-1252, doi:10.1097/01.AOG.0000187893.52488.4b (2005).

17 Cutland, C. L. *et al.* Maternal HIV infection and vertical transmission of pathogenic bacteria. *Pediatrics* **130**, e581-590, doi:10.1542/peds.2011-1548 (2012).

18 Gray, K. J. *et al.* Group B Streptococcus and HIV infection in pregnant women, Malawi, 2008-2010. *Emerg Infect Dis* **17**, 1932-1935, doi:10.3201/eid1710.102008 (2011).

19 Mavenyengwa, R. T. *et al.* Group B Streptococcus colonization during pregnancy and maternal-fetal transmission in Zimbabwe. *Acta Obstet Gynecol Scand* **89**, 250-255 (2010).

20 Madzivhandila, M. *et al.* Serotype distribution and invasive potential of group B streptococcus isolates causing disease in infants and colonizing maternal-newborn dyads. *PloS One* **6**, e17861, doi:10.1371/journal.pone.0017861 (2011).

21 Moyo, S. R., Mudzori, J., Tswana, S. A. & Maeland, J. A. Prevalence, capsular type distribution, anthropometric and obstetric factors of group B Streptococcus (Streptococcus agalactiae) colonization in pregnancy. *Cent Afr J Med* **46**, 115-120 (2000).

22 Mosabi, J. M., Arimi, S. M. & Kang'ethe, E. K. Isolation and characterization of group B streptococci from human and bovine sources within and around Nairobi. *Epidemiology and infection* **118**, 215-220 (1997).

23 Edmond, K. M. *et al.* Group B streptococcal disease in infants aged younger than 3 months: systematic review and meta-analysis. *Lancet* **379**, 547-556, doi:10.1016/S0140-6736(11)61651-6 (2012).

24 The WHO Young Infants Study Group. Bacterial etiology of serious infections in young infants in developing countries: results of a multicentre study. *Pediatr Infect Dis* **18**, S17-22 (1999).

25 Hamer, D. H. *et al.* Etiology of bacteremia in young infants in six countries. *Pediatr Infect Dis J* **34**, e1-8, doi:10.1097/INF.0000000000000549 (2015).

26 Berkley, J., Lowe BS, Mwangi I, Williams T, Bauni E, Mwarumba S, Ngetsa C, Slack MPE, Njenga S, Hart CA, Maitland K, English M, Marsh K, Scott JAG. Bacteremia among Children Admitted to a Rural Hospital in Kenya. *NEJM* **352**, 39 - 47 (2005).

27 Gray, K., Bennett SL, French N, Phiri AJ, Graham SM. Invasive Group B Streptococcal Infections in infants, Malawi. *Emerging Infectious Diseases* **13**, 223 - 229 (2007).

28 Ojukwu JU, A. L., Ugwu J, Orji IK. Neonatal septicemia in high risk babies in South-Eastern Nigeria. *J Perinat Med* **34**, 166-172 (2006).

29 Madhi, S., Radebe K, Crewe-Brown H, Frasch CE, Arakere G, Mokhachane M, Kimura A. High burden of invasive Streptococcus agalactiae in South African Infants. *Annals of Tropical Paediatrics* **23**, 15 - 23 (2003).

30 Folgosa, E. *et al.* A case control study of chorioamniotic infection and histological chorioamnionitis in stillbirth. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* **105**, 329-336 (1997).

31 Moyo, S. R. *et al.* Stillbirths and intrauterine infection, histologic chorioamnionitis and microbiological findings. *Int J Gynaecol Obstet* **54**, 115-123 (1996).

32 Da Cunha, V. *et al.* Streptococcus agalactiae clones infecting humans were selected and fixed through the extensive use of tetracycline. *Nature communications* **5**, 4544, doi:10.1038/ncomms5544 (2014).

33 Sorensen, U. B., Poulsen, K., Ghezzo, C., Margarit, I. & Kilian, M. Emergence and global dissemination of host-specific Streptococcus agalactiae clones. *mBio* **1**, doi:10.1128/mBio.00178-10 (2010).

34 Bekker, V., Bijlsma, M. W., van de Beek, D., Kuijpers, T. W. & van der Ende, A. Incidence of invasive group B streptococcal disease and pathogen genotype distribution in newborn babies in the Netherlands over 25 years: a nationwide surveillance study. *Lancet Infect Dis* **14**, 1083-1089, doi:10.1016/S1473-3099(14)70919-3 (2014).

35 Brochet, M., Couve, E., Bercion, R., Sire, J. M. & Glaser, P. Population structure of human isolates of Streptococcus agalactiae from Dakar and Bangui. *J Clin Microbiol* **47**, 800-803, doi:10.1128/JCM.01103-08 (2009).

36 Mahmmod, Y. S., Klaas, I. C., Katholm, J., Lutton, M. & Zadoks, R. N. Molecular epidemiology and strain-specific characteristics of Streptococcus agalactiae at the herd and cow level. *Journal of dairy science* **98**, 6913-6924, doi:10.3168/jds.2015-9397 (2015).

37 Rato, M. G. *et al.* Antimicrobial resistance and molecular epidemiology of streptococci from bovine mastitis. *Veterinary microbiology* **161**, 286-294, doi:10.1016/j.vetmic.2012.07.043 (2013).

38 Springman, A. C. *et al.* Selection, recombination, and virulence gene diversity among group B streptococcal genotypes. *Journal of bacteriology* **191**, 5419-5427, doi:10.1128/JB.00369-09 (2009).

39 Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4680-4687, doi:10.1073/pnas.1002611107 (2011).

40 Cutland, C. L. *et al.* Increased risk for group B Streptococcus sepsis in young infants exposed to HIV, Soweto, South Africa, 2004-2008(1). *Emerg Infect Dis* **21**, 638-645, doi:10.3201/eid2104.141562 (2015).

41 Le Doare, K. *et al.* Anti-group B Streptococcus antibody in infants born to mothers with human immunodeficiency virus (HIV) infection. *Vaccine* **33**, 621-627, doi:10.1016/j.vaccine.2014.12.025 (2015).

42 Dangor, Z. *et al.* HIV-1 Is Associated With Lower Group B Streptococcus Capsular and Surface-Protein IgG Antibody Levels and Reduced Transplacental Antibody Transfer in Pregnant Women. *J Infect Dis* **212**, 453-462, doi:10.1093/infdis/jiv064 (2015).

43 Johri, A. K. *et al.* Group B Streptococcus: global incidence and vaccine development. *Nature reviews. Microbiology* **4**, 932-942, doi:10.1038/nrmicro1552 (2006).

44 D'Urzo, N. *et al.* Acidic pH strongly enhances in vitro biofilm formation by a subset of hypervirulent ST-17 Streptococcus agalactiae strains. *Appl Environ Microbiol* **80**, 2176-2185, doi:10.1128/AEM.03627-13 (2014).

45 Manning, S., Springman AC, Million AD, Milton NR, McNamara SE, Somsel PA, Bartlett P, Davies HD. Association of Group B Streptococcus colonization and bovine exposure: a prospective cross-sectional cohort study. *PloS One* **5**, e8795 (2010).

46 Huber, C. A., McOdimba, F., Pflueger, V., Daubenberger, C. A. & Revathi, G. Characterization of invasive and colonizing isolates of Streptococcus agalactiae in East African adults. *J Clin Microbiol* **49**, 3652-3655, doi:10.1128/JCM.01288-11 (2011).

47 Almeida, A. *et al.* Whole-Genome Comparison Uncovers Genomic Mutations between Group B Streptococci Sampled from Infected Newborns and Their Mothers. *Journal of bacteriology* **197**, 3354-3366, doi:10.1128/JB.00429-15 (2015).

48 Le Doare, K. *et al.* Risk factors for Group B Streptococcus colonisation and disease in Gambian women and their infants. *J Infect*, doi:10.1016/j.jinf.2015.12.014 (2016).

49 Nasri, K., Chehrei, A. & Manavi, M. S. Evaluation of vaginal group B streptococcal culture results after digital vaginal examination and its pattern of antibiotic resistance in pregnant women. *Iranian journal of reproductive medicine* **11**, 999-1004 (2013).

50 Bellais, S. *et al.* Capsular switching in group B Streptococcus CC17 hypervirulent clone: a future challenge for polysaccharide vaccine development. *J Infect Dis* **206**, 1745-1752, doi:10.1093/infdis/jis605 (2012).

51 Scott, J. A. *et al.* Profile: The Kilifi Health and Demographic Surveillance System (KHDSS). *Int J Epidemiol* **41**, 650-657, doi:10.1093/ije/dys062 (2012).

52 Schrag, S. J. & Verani, J. R. Intrapartum antibiotic prophylaxis for the prevention of perinatal group B streptococcal disease: Experience in the United States and implications for a potential group B streptococcal vaccine. *Vaccine*, doi:10.1016/j.vaccine.2012.11.056 (2012).

53 Stoner, K., Rabe LK, Hillier SL. Effect of Transport Time, Temperature, and Concentration on the Survival of Group B Streptococci in Amies Transport Medium. *J Clin Microbiol* **42**, 5385-5387 (2004).

54 Mathers, A. J. *et al.* Klebsiella pneumoniae carbapenemase (KPC)-producing K. pneumoniae at a single institution: insights into endemicity from whole-genome sequencing. *Antimicrob Agents Chemother* **59**, 1656-1663, doi:10.1128/AAC.04292-14 (2015).

55 Sheppard, A. E. *et al.* Capsular typing method for Streptococcus agalactiae using whole genome sequence data. *J Clin Microbiol*, doi:10.1128/JCM.03142-15 (2016).

56 Didelot, X. & Wilson, D. J. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol.* **11** (2015).

57 Filmer, D. & Pritchett, L. H. Estimating wealth effects without expenditure data--or tears: an application to educational enrollments in states of India. *Demography* **38**, 115-132 (2001).

Acknowledgements

We thank The Wellcome Trust [093804] for funding this study. ACS and JAB are funded by fellowships from the Wellcome Trust (www.wellcome.ac.uk [093804, 098532]). DWC is a NIHR (UK) Senior Investigator. DWC, AES and ASW received funding from the Health Innovation Challenge Fund (a parallel funding partnership between the Department of Health and the Wellcome Trust) [HICF-T5-358 and WT098615/Z/12/Z]; the UK Clinical Research Collaboration (a parallel funding partnership between the Medical Research Council [G0800778], Biotechnology and Biological Sciences Research Council and the Wellcome Trust [087646/Z/08/Z]); and the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre. SS receives funding from the National Center for Immunization and Respiratory Diseases, US Centers for Disease Control and Prevention. The Wellcome Trust (core grant [077092]) and the Bill and Melinda Gates Foundation fund paediatric and maternal research at KEMRI-Wellcome Trust Programme.

We also thank all the fieldworkers and clinical staff who contributed to this work in Kilifi County Hospital (formerly Kilifi District Hospital), Coast Provincial General Hospital, Bamba sub-district Hospital and Ganze Health Facility, as well as all the participants in this study. Whole genome sequencing was undertaken at The Wellcome Trust Centre for Human Genetics, University of Oxford and we thank the library and sequencing teams. Initial pre-processing of raw sequence data was done using a data processing pipeline developed by the Department of Statistics, University of Oxford. Surveillance at Kilifi County Hospital was undertaken at the Kenya Medical Research Institute/Wellcome Trust Research Programme and we thank all those involved. This study is published with the permission of the Director of Kenya Medical Research Institute.

Contributions

The study was conceived and designed by ACS, ACK, SCM, CJ, BT, SJS, SHK, GD, DWC, and JAB. Data were acquired, analysed and/or interpreted by ACS, ACK, AES, HCB, JL, EA, SM, SM, KA, AV, AG, PM, LW, HM, DM, MS, BK, NM, EM, DM, VB, MS, O, NO, ASW, SJS, GF, DWC, JAB. Administrative or technical support was given by AES, SM, SCM, KA, AV, AG, PM, LW, CJ, NM, BT, EM, DM, VB, MS, MO, NO, ASW, SHK, GF, DWC, and JAB. Statistical analysis was done by ACS, with advice from GF, ASW and JAB. Phylogenetics were done by AES with ACS. The first draft was written by ACS. All authors reviewed the manuscript.

Competing financial interests

We declare no competing interests.

Corresponding author

Correspondence to Anna Seale KEMRI-Wellcome Trust, PO Box 230, Kilifi 80108, Kenya. +254 41 7522155 aseale@nhs.net

Table 1: Exposures associated with maternal Group B Streptococcus (GBS) colonisation

		No GBS[c]	GBS[c]		Complete cases (N=3979)			Imput
Variable		**N not missing**	**N not missing**	**(%)**	**OR**	**95%CI[a]**	**p[b]**	**OR**
Site	Rural	479	47	(8.9)	0.80	(0.73-0.88)	<0.001	0.91
	Semi-rural	4862	608	(11.1)	1			1
	Urban	1692	279	(14.2)	0.96	(0.93-1.00)		0.95
Age in quartiles (years)[e]	<21.5	1674	166	(9.0)	0.77	(0.55-1.15)	0.009	0.80
	21.5-25.3	1663	223	(11.8)	1.15	(0.87-1.22)		1.03
	25.4-29.9	1656	213	(11.4)	1			1
	≥30	1672	186	(10.0)	0.91	(0.78-1.18)		0.96
Parity	0	2986	365	(10.9)	1.06	(0.99-1.09)	<0.001	1.05
	1-4	3550	442	(11.1)	1			1
	≥5	1341	119	(8.2)	0.85	(0.69-0.92)		0.81
Ethnicity: Mijikenda[d]	No	2226	345	(13.4)	1		0.002	1
	Yes	5617	578	(9.3)	0.65	(0.60-0.90)		0.73
Household socioeconomic status (quartiles)[e]	Very low	1086	96	(8.1)	0.88	(0.66-1.16)	<0.001	0.89
	Low	2720	294	(9.8)	1			1
	Medium	2123	229	(9.7)	1.00	(0.82-0.92)		0.88
	High	2038	315	(13.4)	1.24	(1.06-1.30)		1.21
Mother looks after cattle	No	7471	873	(10.5)	1		<0.001	1
	Yes	449	56	(11.1)	1.46	(1.17-1.42)		1.29
Nutritional status (mid-upper arm circumference in cm[e])	≤23.9	1428	125	(8.0)	0.77	(0.60-0.89)	<0.001	0.72
	24-25.9	2219	264	(10.6)	1			1
	26-27.9	1662	183	(9.9)	0.80	(0.66-1.07)		0.85
	≥28	2170	309	(12.5)	1.02	(0.78-1.40)		1.05
HIV infection	No	7285	879	(10.8)	1		<0.001	1
	Yes, no CTX[f]	239	20	(7.7)	1.16	(0.92-1.45)		0.68
	Yes, on CTX[f]	161	5	(3.0)	0.20	(0.14-0.26)		0.24
Vaginal examination before swab	No	4952	609	(11.0)	1		0.019	1
	Yes	780	73	(8.6)	0.57	(0.36-0.91)		0.83
Obstetric complication	No	6913	823	(10.6)	1		<0.001	1
	Yes	1054	111	(9.5)	0.78	(0.70-0.88)		0.85

[a] 95% confidence intervals are given, based on robust standard errors to account for intracluster correlation within recruitment sites
[b] p values are derived from the Wald test (imputations combined using Rubin's rules)
[c] Full details on all variables and numbers for missing variables are given in Supplementary Table 4
[d] Mijikenda are the indigenous coastal population
[e] For continuous variables we tested for associations prior t
non-linearity, natural cubic splines were used (Supplementa
the largest group was used as the reference group.
[f] CTX=co-trimoxazole prophylaxis

Table 2: Exposures associated with maternal Group B Streptococcus (GBS) colonisation wit

		GBS			Univariable complete cases (N=914)			Mu (N=
Variable		**Not CC17**	**N CC17**	**(%)**	**OR**	**95%CI[a]**	**p[b]**	**OR**
Site	Rural	33	13	28.3	0.85	(0.43-1.65)	0.072	
	Semi-rural	403	187	31.7	1			
	Urban	211	67	24.1	0.68	(0.49-0.95)		
Age in quartiles (years)[d]	<21.5	115	49	29.9	1.16	(1.07-1.27)	0.2	
	21.5-25.3	156	60	27.8	1.05	(0.92-1.21)		
	25.4-29.9	153	56	26.8	1			
	≥30	130	51	28.2	1.07	(0.85-1.35)		
Parity	0	257	98	27.6	0.86	(0.49-1.51)	0.4	
	1 to 5	301	133	30.6	1			
	≥5	83	35	29.7	0.95	(0.83-1.10)		
Ethnicity: Mijikenda[c]	No	262	79	23.2	1		<0.001	
	Yes	379	183	32.6	1.60	(1.52-1.69)		
Household socioeconomic status[d] (quartiles)	Very low	71	25	26.0	0.61	(0.48-0.80)	<0.001	
	Low	192	95	33.1	1			
	Medium	155	69	30.8	1.21	(0.82-1.80)		
	High	229	78	25.4	0.69	(0.41-1.15)		
Mother looks after cattle	No	598	255	29.9	1		<0.001	
	Yes	44	12	21.4	0.64	(0.58-0.70)		
Nutritional status (mid-upper arm circumference in cm[d])	≤23.9	81	41	33.6	1.19	(0.57-2.56)	0.0042	
	24-25.9	181	77	29.8	1			
	26-27.9	130	48	27.0	0.87	(0.56-1.35)		
	≥28	219	85	28.0	0.91	(0.57-1.46)		
HIV infection	No	608	251	29.2	1		<0.001	
	Yes, no CTX[e]	13	7	35.0	1.30	(1.21-1.40)		
	Yes, on CTX[e]	2	3	60.0	3.63	(1.58-8.34)		

[a] 95% confidence intervals are given, based on robust standard errors to account for intracluster correlation within recruitment sites
[b] p values are derived from the Wald test
[c] Mijikenda are the indigenous coastal population
[d] For continuous variables we tested for associations prior to categorisation and inclusion in the model. Where there was non-linearity, natural cubic splines were used (Supplementary Figure 2). Data were categorised for ease of presentation, and the largest group was used as the reference group.
[e] CTX=co-trimoxazole prophylaxis

Figure 1: Study design and recruitment of participants by study site
a, Recruitment timeline and sub-studies undertaken at each study site. **b,** Recruitment of mothers in the cohort study. *The denominator for live-births in the prospective cohort period, used to calculate incidence of early onset disease in Kilifi County Hospital (KCH) excluded those who did not deliver, or had a stillbirth (leaving 6598.**These mothers (7967) were included in the analysis of risk factors for maternal GBS colonisation. [§]These births (7833) were included in analyses assessing GBS as a risk factor for stillbirth or perinatal death. [§§]These live-births (7408) were included in analyses assessing GBS as a risk factor for preterm birth, low birth-weight or possible serious bacterial infection. **c** Recruitment for the vertical transmission study (maternal-neonatal dyads), a subset of mothers who delivered in KCH. **d** Recruitment for stillbirth nested case-control study including mothers who delivered in KCH and had a stillbirth, and controls.

Figure 2: Interaction of risk factors at delivery with maternal GBS colonisation associated with adverse newborn outcomes.

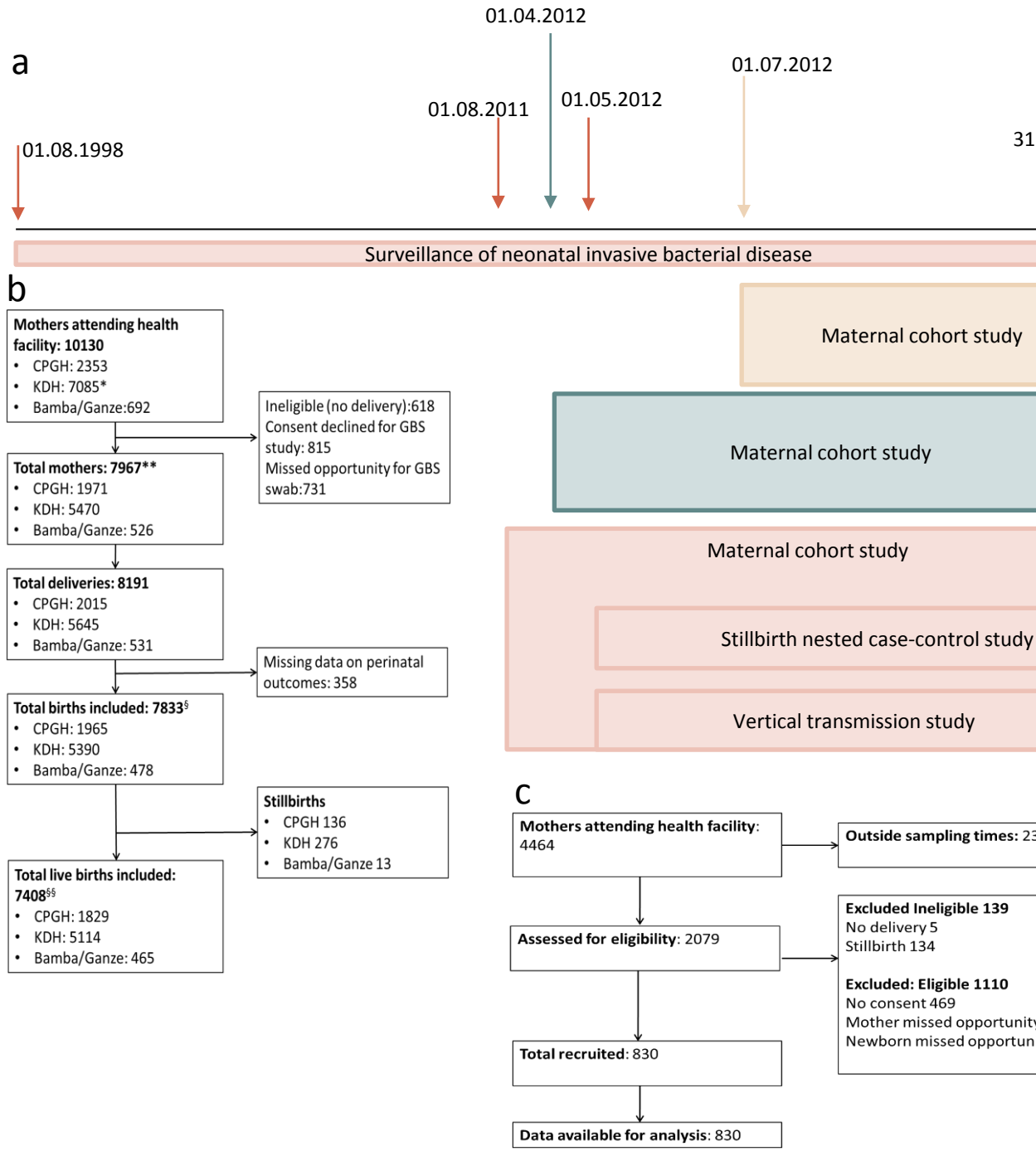
Interactions between maternal risk factors at delivery (maternal fever, maternal urinary tract infection, prolonged rupture of membranes) and adverse perinatal outcomes (very preterm birth, very low birth weight, stillbirth, possible serious bacterial infection), in the presence and absence of maternal GBS colonisation. Odds ratios are given for maternal exposures and associated perinatal outcome (listed vertically) with 95% confidence intervals illustrated with error bars for the odds ratio in each case. Interactions were included in multivariable models if there was evidence of interaction at the $p<0.1$ level in univariable analyses. P values given here are for interaction tests in imputed multivariable models (details for all models in Supplementary tables 5-9). **Possible serious bacterial infection (pSBI) is defined in Supplementary table 1; it is a clinical diagnosis used to guide empiric treatment of neonates for possible serious bacterial infections in resource-poor settings.

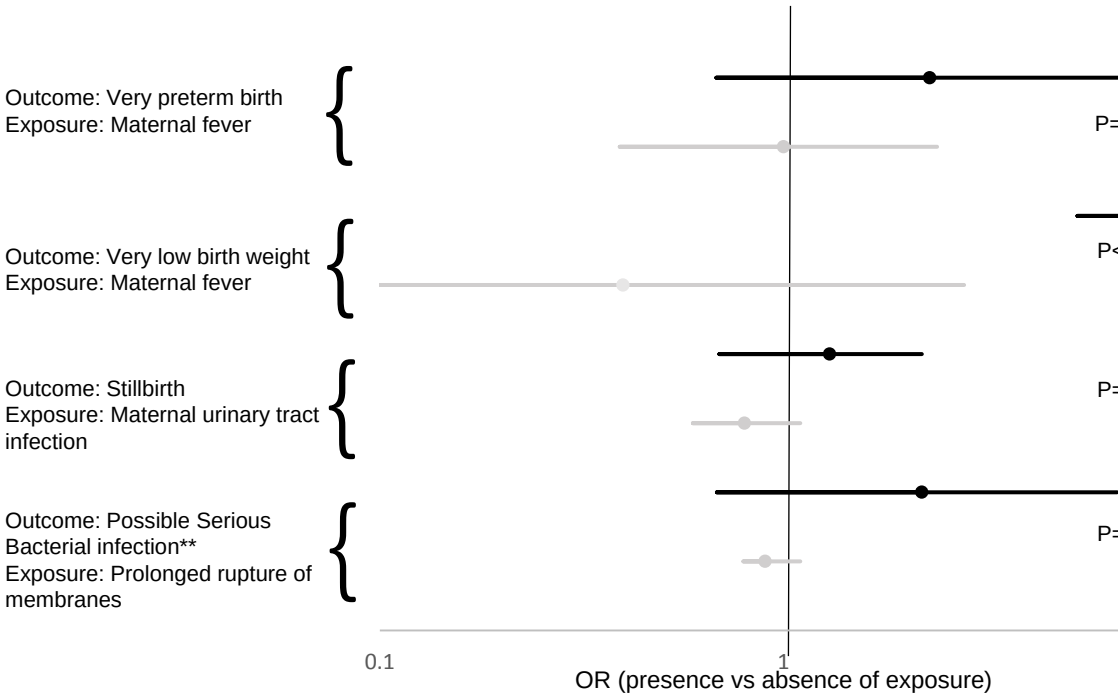
Figure 3: GBS types colonising mothers and causing disease.

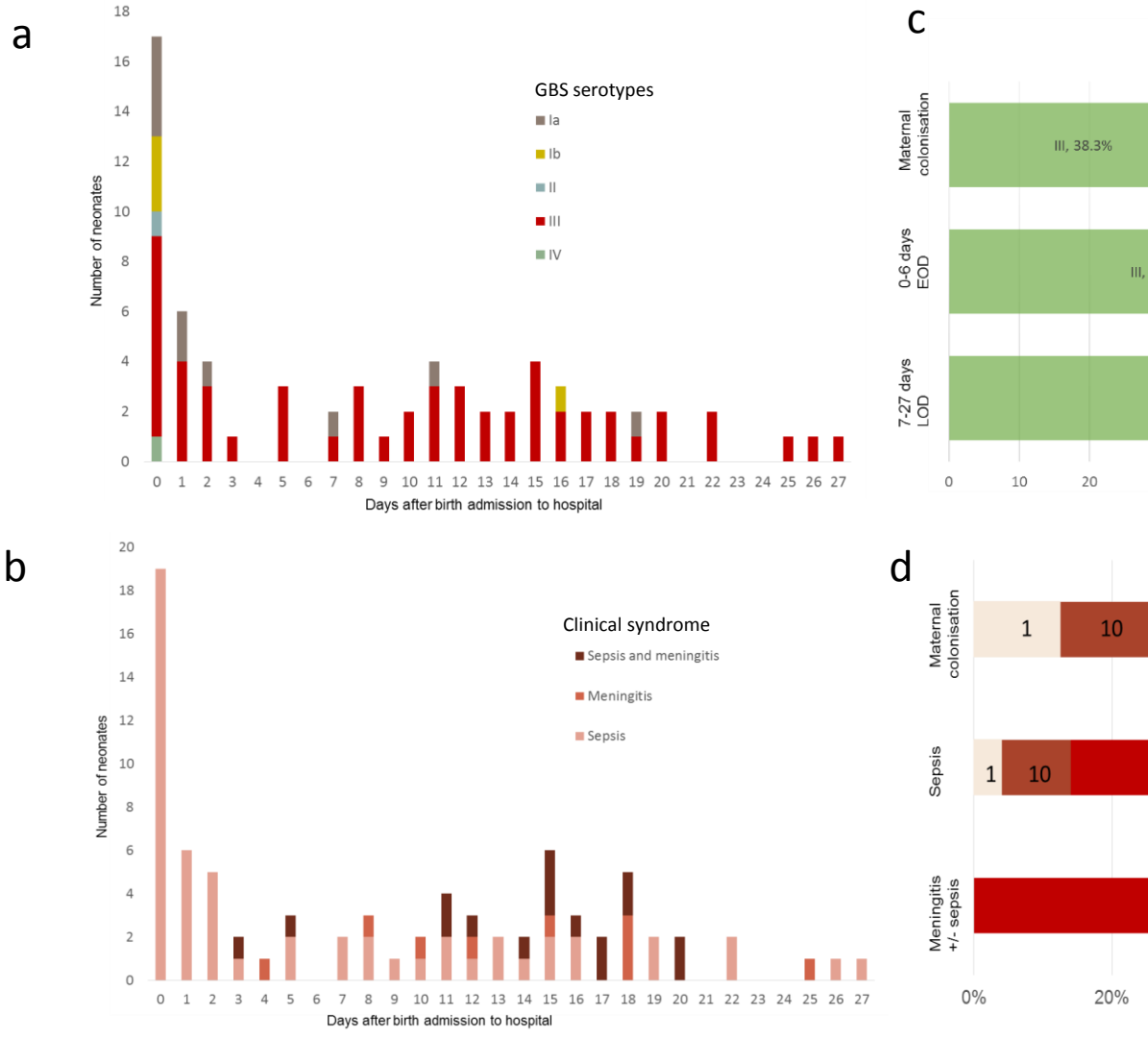
a, Invasive neonatal GBS disease cases decrease after the first few days of birth in Kilifi County Hospital neonatal admissions (1998-2013), and serotype III causes an increasing proportion of disease; **b**: The clinical infection syndrome is predominantly sepsis in the first few days after birth in neonates admitted with invasive GBS disease to Kilifi County Hospital (1998-2013) with increasing numbers of neonates admitted with meningitis with or without sepsis later in the neonatal period; **c**, The percentage of different serotypes in GBS isolates from maternal colonisation, early onset disease (EOD) and late onset disease (LOD) in neonates shows a stepwise increase in serotype III from maternal colonisation to EOD and LOD; **d**, The percentage of different clonal complexes in GBS isolates from maternal colonisation, neonatal sepsis and neonatal meningitis (+/- sepsis) shows the increasing dominance of CC-17 in neonatal disease, particularly in neonatal meningitis.

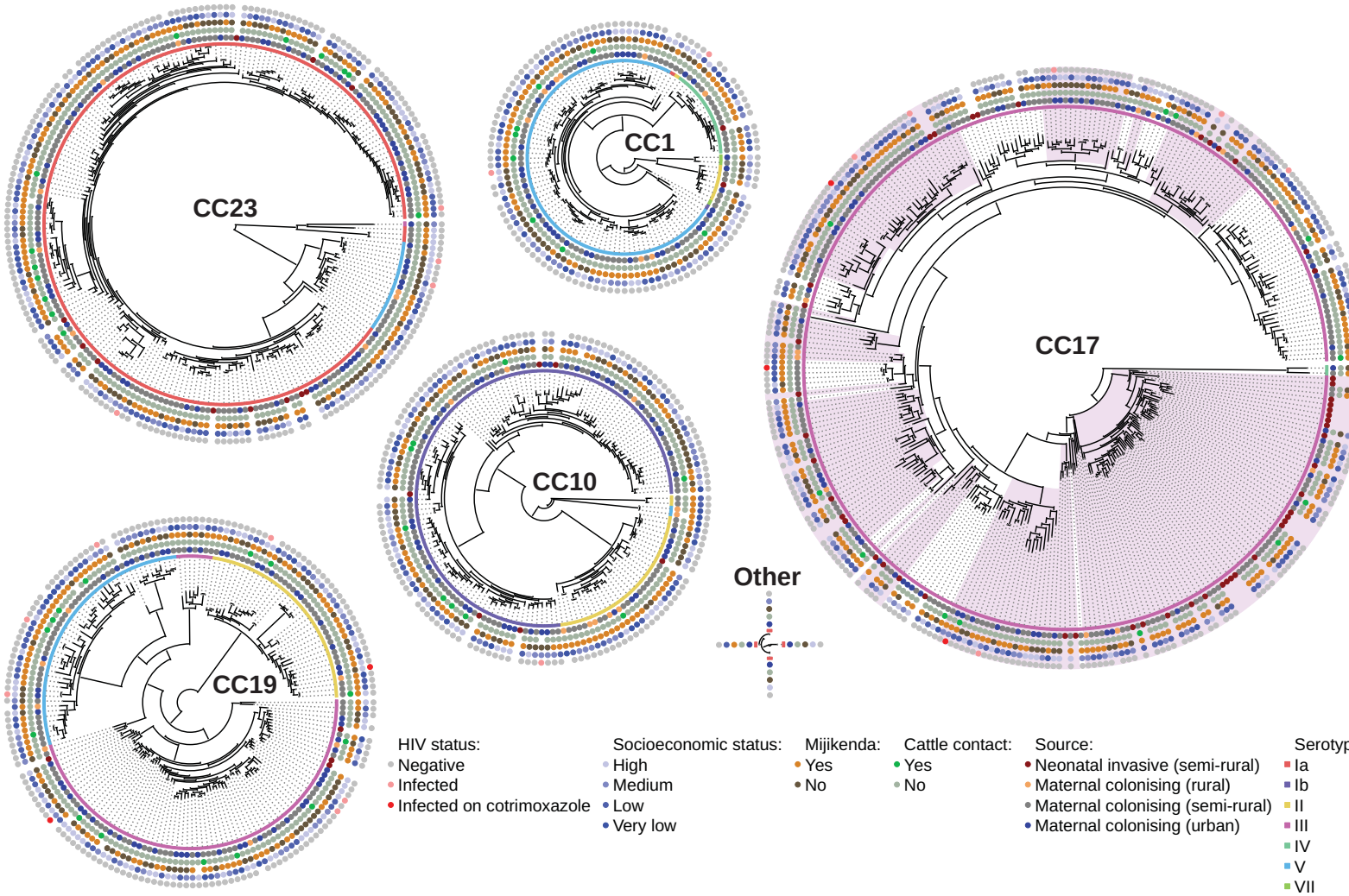
Figure 4: Phylogenetic reconstructions of GBS isolates

Maximum likelihood phylogenies, with recombinant regions removed, are shown separately for each clonal complex. Background shading indicates ST-17 isolates within CC-17. Serotypes are illustrated for each clonal complex in the innermost circle. The next circle describes the sample source of the GBS isolate (neonatal invasive, or maternal colonising (by site of recruitment)). For maternal colonising isolates, epidemiological details are illustrated. From the outermost circle, these are: maternal HIV status (negative, HIV-infected, HIV infected and taking prophylactic co-trimoxazole), socio-economic status (high, medium, low and very low), ethnicity (Mijikenda or non-Mijikenda) and the presence or absence of cattle contact.


a
01.04.2012
01.07.2012
01.08.2011
01.05.2012
01.08.1998
31
Surveillance of neonatal invasive bacterial disease
b
Mothers attending health facility: 10130
• CPGH: 2353
• KDH: 7085*
• Bamba/Ganze:692
Ineligible (no delivery):618
Consent declined for GBS study: 815
Missed opportunity for GBS swab:731
Total mothers: 7967**
• CPGH: 1971
• KDH: 5470
• Bamba/Ganze: 526
Total deliveries: 8191
• CPGH: 2015
• KDH: 5645
• Bamba/Ganze: 531
Missing data on perinatal outcomes: 358
Total births included: 7833§
• CPGH: 1965
• KDH: 5390
• Bamba/Ganze: 478
Stillbirths
• CPGH 136
• KDH 276
• Bamba/Ganze 13
Total live births included: 7408§§
• CPGH: 1829
• KDH: 5114
• Bamba/Ganze: 465
Maternal cohort study
Maternal cohort study
Maternal cohort study
Stillbirth nested case-control study
Vertical transmission study
c
Mothers attending health facility: 4464
Outside sampling times: 23
Assessed for eligibility: 2079
Excluded Ineligible 139
No delivery 5
Stillbirth 134
Excluded: Eligible 1110
No consent 469
Mother missed opportunity
Newborn missed opportun
Total recruited: 830
Data available for analysis: 830


Outcome: Very preterm birth
Exposure: Maternal fever
Outcome: Very low birth weight
Exposure: Maternal fever
Outcome: Stillbirth
Exposure: Maternal urinary tract infection
Outcome: Possible Serious Bacterial infection**
Exposure: Prolonged rupture of membranes
P=
P<
P=
P=
0.1
1
OR (presence vs absence of exposure)


a
b
c
d
GBS serotypes
Ia
Ib
II
III
IV
Number of neonates
Days after birth admission to hospital
Clinical syndrome
Sepsis and meningitis
Meningitis
Sepsis
Maternal colonisation
0-6 days EOD
7-27 days LOD
III, 38.3%
Sepsis
Meningitis +/- sepsis
1
10
0%
20%


CC23
CC1
CC17
CC10
Other
CC19
HIV status:
Negative
Infected
Infected on cotrimoxazole
Socioeconomic status:
High
Medium
Low
Very low
Mijikenda:
Yes
No
Cattle contact:
Yes
No
Source:
Neonatal invasive (semi-rural)
Maternal colonising (rural)
Maternal colonising (semi-rural)
Maternal colonising (urban)
Serotyp
Ia
Ib
II
III
IV
V
VII