(2700 Characters including authors list/affiliations with spaces)

Title: Large-scale data analysis on endometriosis from the UK Biobank reveals novel genetic loci

Authors: Nilufer Rahmioglu[1,2], James Cook[3], International Endometriosis Genomic Consortium (IEGC), Christian Becker[2], Andrew Morris[1,3], Krina Zondervan[1,2].

Affiliations: [1]Wellcome Center for Human Genetics, Oxford, UK; [2]Endometriosis CaRe Centre Oxford, Nuffield Dept of Obstetrics and Gynaecology, Oxford University, UK; [3]Dept of Biostatistics, Liverpool University, UK.

Introduction: Endometriosis is a common complex condition with an estimated heritability of 50%. The largest meta-analysis of genome-wide association studies (GWAS) for endometriosis to date, including 17K cases, identified 14 genome-wide significant loci ($p<5x10^{-8}$) together explaining only 5% of the heritability. In ongoing analyses we are substantially increasing sample size by including novel datasets such as the UK Biobank. The UK Biobank involves medical records and self-reported questionnaire data on 500,000 volunteer participants, for which the full genotype data, imputed up to the HRC reference panel, has been recently released.

Methods: Of the 274K white British women participating in the UK Biobank, there are 8,187 endometriosis cases, of which 4,196 are self-reported (90% with a related diagnostic surgery) and 5,024 have a relevant ICD9/10 code. We have conducted genome-wide association testing of variants with minor allele frequency >1% using a linear mixed model (LMM) to account for relatedness and population structure.

Results: We replicated all 14 known endometriosis loci ($p<0.05$, same direction of effect). We also identified two novel associations at genome-wide significance. The first maps to *IGF2BP3* (rs550554633, $p=4.4x10^{-8}$), which is involved in regulation of the metabolism and adiposity, supporting the genetic sharing between endometriosis and fat distribution. The second maps to *GDAP1* (rs554964149, $p=7.2x10^{-8}$), which is involved in neuronal development that may play a role in endometriosis-associated pain, and previously has been associated with pain severity in dysmenorrhea. Moreover, 18 loci near genome-wide significance ($p<5x10^{-6}$) were identified, which require validation in additional datasets.

Conclusions: Ongoing consortium analyses include meta-analysis of UK Biobank results with other novel and existing datasets (>30K cases), which is highly likely to reveal further novel associations that will be presented at the meeting. These will shed light on the underlying genetic mechanisms of this debilitating condition and may point to novel therapeutic interventions.

Keywords: Endometriosis, GWAS, UK Biobank.