**Systems-Level Insights into Bronchopulmonary Dysplasia from Meta-Analysis of Genome-Scale Studies**

Chris Happs1,2\*, Prerna Khanna1,2\*, Charlotte Scoynes2, Nelly Muriungi1, Akira Alexander1, Josh Rogers1, Elizabeth Lake1,2, Maximillian Pickup1,2, Nicholas Parkinson1, Jonathan Millar2, Kenneth Baillie1,2, Sara Clohisey Hendry1,2

\* These authors contributed equally

1. Roslin Institute, University of Edinburgh, Edinburgh, U.K.
2. Baillie-Gifford Pandemic Science Hub, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, U.K.

**Running title:** Meta-analysis of Bronchopulmonary Dysplasia

**ABSTRACT**

**Background**

Despite marked improvements in survival following preterm birth, the incidence of bronchopulmonary dysplasia (BPD) continues to rise. BPD remains the most prevalent complication of prematurity and carries the risk of long-term morbidity. Characterising the cellular and molecular mechanisms driving disease progression is critical for informing clinical management and improving outcomes. To this end, we conducted a meta-analysis of genome-scale studies to identify molecular pathways implicated in BPD progression in both human cohorts and animal models.

**Methods**

Gene lists associated with bronchopulmonary dysplasia in humans and in rodent models were obtained from systematically identified genome-scale studies. These lists were then analysed using the meta-analysis by information content (MAIC) algorithm, which integrates multiple datasets to produce a single aggregated, ranked gene list based on the cumulative strength of evidence for each gene.

**Results**

**Conclusions**

**Key words:** Bronchopulmonary Dysplasia; adaptive immunity, T cells,

**Background**

Bronchopulmonary dysplasia (BPD), also known as Chronic Lung Disease, is one of the most common complications of preterm birth, affecting up to 30% of preterm infants born before 32 weeks post-menstrual age (PMA).(*Neonatal Data Analysis Unit (NDAU). Neonatal Health Intelligence Tool. 2021. Available: Https://Www.Imperial.Ac.Uk/Neonatal-Data-Analysis-Unit/Neonatal-Data-Analysis-Unit/Neonatal-Data-Visualisations/*, n.d.),(Isayama et al., 2017; Jensen et al., 2019) Antenatal and postnatal factors disrupt the developing lung leading to decreased alveolarisation, larger alveoli, irregular pulmonary vessels, and fibrotic tissue.(Gilfillan et al., 2021; Thébaud et al., 2019) The effects of BPD are life-long; individuals with BPD are more likely to require rehospitalisation during their childhood, to have limited lung function and to display delayed neurodevelopment, compared to those born preterm without BPD.(Dassios & Greenough, 2021; Greenough, 2000, 2006; Sun et al., 2023)

Understanding the molecular mechanisms that contribute to BPD could aid in risk prediction, targeted therapies, and understanding disease mechanisms. The heterogeneity of the disease has hampered efforts to identify reliable, consistent biomarkers. Alongside evidence from twin studies pointing to a possible heritable component in BPD(Bhandari et al., 2006; Lavoie et al., 2008) some candidates have been identified (e.g. SPOCK2, CRP). Though no gene has been significantly associated with the development of BPD via Genome Wide Association Studies (GWAS) or exome sequencing.(Ambalavanan et al., 2015; Hadchouel et al., 2011; Mahlman et al., 2017; Torgerson et al., 2018; H. Wang et al., 2013) Recently, transcriptomic data has been used to identify four endotypes of BPD, distinguished by T helper cell and T cell signalling.(Moreira et al., 2023) Integrating results from heterogenous sources can allow for novel insights, reflecting an ongoing effort to incorporate insights from genome-scale research, complementing the longstanding emphasis on clinical studies in the field.

Comparative transcriptomic analyses across species can provide valuable insights into common mechanisms while also highlighting potential species-specific responses. Rodent models, particularly murine models, have been widely employed to study the pathophysiology of BPD. Anatomical and physiological differences have prevented in-depth research of ventilator induced lung injury in this model and research has concentrated primarily on hyperoxia-induced lung injury, mimicking the impaired alveolar development and inflammation observed in preterm infants.(Hurskainen et al., 2021)

We conducted two concurrent systematic reviews of the BPD literature, focusing on genome-scale studies in humans and rodents. Results were extracted from identified studies, where possible, and analysed using Meta Analysis by Information Content (MAIC) algorithm. MAIC is a previously described algorithm that uses data-driven gene list weightings to produce a comprehensive ranked list of genes associated with the trait of interest.(Millar et al., 2024; Parkinson et al., 2020; The GenOMICC Investigators et al., 2021; B. Wang et al., 2022)

Our findings demonstrate that BPD is a complex disease which involves dysregulation in adaptive immune signalling, and extracellular matrix organisation. Of note, T cell-associated genes feature prominently suggesting, with recent studies, that adaptive immunity may play a larger role in BPD pathogenesis, and potentially ventilator associated lung injury, than previously appreciated.

**Results**

*Systematic review*

We conducted a systematic review of studies that reported associations between genes, transcripts, or proteins, and the development of BPD in humans or the study of BPD in animal models. We did not limit the search to any specific definition of BPD in humans or animal models. Our search yielded 4450 unique citations that were evaluated for inclusion [Supplementary Figures 1 and 2]. For full text evaluation we retrieved 76 articles for human studies and 118 for animal studies, without overlap. Although large animal models are used to explore BPD,(Eiby et al., 2013) we found none that met our inclusion criteria (Supplementary Tables 1 and 2) and the analysis therefore focused on rodents(mice and rats).

[references for Table 1(Ahmed et al., 2023; Ambalavanan et al., 2015; Bhattacharya et al., 2012, 2020)]

*Meta-Analysis by Information Content of human studies*

From the 81 human studies, 21 met our eligibility criteria[Supplementary Figure 1, Table 1]. This yielded 31 gene lists, 23 ranked and 8 unranked, employing 5 experimental techniques (Supplementary Table 3,) and representing 4358 infants (cases = 2247, controls = 2111). While the definition of BPD varied between studies, over two-thirds of studies, (71.4%, 15/21), used receipt of respiratory support at 36 weeks PMA.(Jensen et al., 2019) All but a single study used primary tissue or samples. MAIC collated and ranked 8017 unique genes across the datasets [Supplmentary Table 4, Figure1A], and only a single gene was supported by evidence from all 5 experimental techniques (*RASGRP3*, rank = 1). Few genes (n = 53 genes, 0.6%) were supported by >=3 methods while 745 genes (9.74%) were supported by more than 1 method. There is a focus on immunity (*RASGRP3, IL1R2, CD2, CHIT1, GPA33, CD177, RCAN3)* and structure *(LRRN3, FBN1)* within the top 10 ranked genes. Additionally, *HGB1* features, a subunit of foetal hemoglobin. We prioritised 945 genes for further investigation [Figure 1B], as previously described.(Millar et al., 2024)

*Over representation and enrichment analyses*

We first performed Over Representation Analysis (ORA), on the prioritised list of 945 genes. The ORA results from the Gene Ontology database (encompassing all ontologies) were grouped first based on semantic similarity and then based on function [Figure 1C]. Our analysis reveals a significant over-representation of adaptive immune system processes, with a prominent emphasis on T cell differentiation and regulation, as well as cell-cell adhesion. This is further supported by ORA using the KEGG and Reactome databases [Supplementary Figures]. T cell development was further highlighted when we carried out gene set enrichment analysis (GSEA) of the human BPD MAIC dataset [Figure 1D, E].

We then created a protein-protein interaction (PPI) network using the prioritised set of genes. MCL clustering identified 25 clusters with >=5 members. The 7 largest clusters contained >=10 members [Supplementary Figures shows the top 5 clusters]. Using previously described methods (Millar et al., 2024) we identified 32 hub genes within the PPI network, suggested as being central to the wider network [Figure 1F]. MCL clustering reveals , (Figure 1G).

*ARDS dataset*

Acute lung injury caused or worsened by mechanical ventilation during treatment is a feature of BPD in neonates and Acute Respiratory Distress Syndrome (ARDS) in adults (paediatric ARDS in children). We sought to understand firstly, the different deleterious effects mechanical ventilation may have on the developed versus the developing lung, and to understand the common factors that may underlie both syndromes. To examine this, we evaluated the human BPD MAIC dataset against our previously generated ARDS MAIC analysis.(Millar et al., 2024)

Over Representation Analysis focusing on the prioritised genes from both datasets highlighted shared pathology relating to adaptive immunity with pathways associated with positive regulation of T cell activation [Figure2A]. While NF-kappa signal transduction is associated with ARDS, it does not have a footprint in our BPD data, indicating a divergence in pathologies.

GSEA

We then sought the overlap between both datasets, to find the genes most likely to be associated with both syndromes from the literature. There were 112 genes that overlapped in the prioritised sets from both datasets, [Figure2B] and this overlap was found to be significant using a hypergeometric test using the protein-coding genome as the background set (p<0.005). Over representation analysis found leukocyte mediated immunity, lymphocyte mediated immunity and positive regulation of both T cells and type II interferon were highlighted [Figure2C]. A PPI network focused on the overlap between datasets identifies a hub consisting of 5 genes (CD4, CXCL8, MMP9, CD40 and CD2) [Figure 2D].

*Meta-Analysis by Information Content of rodent studies*

From 20 eligible rodent BPD studies [Supplementary Figure 2, Table 2] we extracted 35 gene lists for MAIC analysis (28 ranked and 7 unranked), employing 5 experimental techniques (Supplementary Table 3). Of these studies, 13 used mice models and 7 used rat models. The study therefore reflects primarily a rodent dataset rather than the broader non‑human mammalian dataset originally intended. The majority (17/20; 85%) of studies modelled BPD by exposing neonatal pups to a hyperoxic environment, though the approaches differed; there was variation in the timeframe of hyperoxia (from 3 days to 14 days), recovery in room air (0 days to 57 days), continuous versus intermittent hyperoxia, and the percent oxygen (70 – 100%) used across the studies. Other methods to model BPD in rodent models included lipopolysaccharide (LPS) treatment (1 study) and induced IUGR through maternal diet (1 study). All but one included study used lung tissue for analyses. MAIC collated 7637 genes across all datasets [Figure 3A, Supplementary table 5] and 3 genes were supported by evidence from >=4 experimental techniques (*RAC2*, *CXCR2*, *IL1R2*). One in five genes were supported by more than 1 method (n = 1597, 20.91%), with 141 genes (1.85%) supported by >=3 methods. Immune processes are again represented within the top ten genes (*RAC2*, *CXCR2*, *IL1R2*, *CCDC3*, *PTGS2*). Interestingly, *IL1R2* is shared between the top 10 lists of both analyses. We prioritised 1783 genes for further investigation [Figure 3B].

*Rodent functional enrichment*

Gene ontology ORA revealed a significant overrepresentation of pathways related to myeloid cells and development of muscle tissue and extracellular matrix organisation[Figure 3C]. Regulation of T cell activation is also observed though is far less prominent than seen in the human results. As before, we created a protein-protein interaction (PPI) network. MCL clustering identified 55 clusters with >=5 members. The 21 largest clusters contained >=10 members. Among these, we found programs associated with the adaptive immune response and mitosis. We used the PPI network to identify 13 hub genes, suggested as being central to the wider network [Figure 3D].

*Comparison of human and rodent datasets*

Ontology ORA showed that while the adaptive immune signal observed in humans above was seen in the rodent dataset, the rodent dataset also highlights the role of the extracellular matrix and wound healing [Figure 4A]. The gross overlap between the two parallel MAIC analyses carried out in this study was not shown to be significant by hypergeometric test (p>0.5). However, focusing on the prioritised genes from both datasets, a hypergeometric test, assuming that 99% of human protein coding genes can be mapped to the mouse genome, indicated the overlap of 163 genes [Figure 4B] was significant [Supplementary Table 6]. It is difficult to determine if some uncommon, enriched pathways reflect rodent-specific responses to experimental injury rather than universally conserved disease mechanisms.

**Discussion**

The clinical and biological heterogeneity of BPD presents challenges for identifying the molecular processes that drive progression of the disease. To address this, we applied a validated *in silico* approach to systematically integrate and prioritise existing genome-scale BPD datasets from both human and rodent studies. We further compared our results to previously generated similar output for ARDS generated within our lab.

IL1R2 is upregulated in acute hypoxia.(Johnson et al., 2007)

**T cells in neonates**

Our findings prioritise specific elements of the adaptive immune response In particular T cell development and adhesion. Interestingly, [subgroup based on T cells]. But does this reflect subgroups of phenotype (<https://doi.org/10.1164/rccm.201907-1342OC>**)**

**T cells in BPD – long term effects?**

Altered populations of T cells in adults following BPD,

This study is based on the published literature and is intended to capture the current state of genome-wide BPD research; as such, it is subject to its limitations. Firstly,

Secondly, blood is among the least invasive and most practical biological samples to collect in the NICU setting. Over half of the gene lists from human studies were derived from primary blood samples (11/21, 52.4%). Consequently, our human results are likely indicative of circulating biomarkers and may not entirely reflect the extent of tissue injury at its origin. Conversely, blood samples are not commonly taken in rodent studies of BPD, where whole lung tissue is preferred. This may have contributed to the different gene signatures observed between the two analyses.

Finally, due to key developmental differences - such as timing of alveolarisation and immune system maturation – the results may not be directly comparable between species.

In conclusion, ...

**Methods**

***Systematic review***

The systematic review and meta-analysis protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO; CRD42022306270, CRD42024550229). The review is reported in compliance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.(Page et al., 2021)

***Search strategy and selection criteria***

A detailed description of our search strategy and eligibility criteria is provided in the Supplementary Methods. Briefly, we searched MEDLINE, Embase without language restrictions on 24/05/2024. We included human genome-wide studies reporting associations between genes, transcripts, or proteins. For human studies we accepted any contemporaneous BPD definition. For rodent studies, definitions of BPD study groups were accepted as hyperoxic exposure and induced low birth weight. We excluded candidate in vivo or in vitro studies (< 50 genes/proteins), candidate gene associations, and studies with <5 patients per arm. Following deduplication, titles were initially screened using Screenatron.(Clark et al., 2020)

Abstracts were then screened against eligibility criteria, this included organising the studies into separate collections of human and animal studies, with an independent author resolving inconsistencies. Full texts were retrieved and analysed for inclusion before extraction of gene lists for inclusion in MAIC. Input lists were processed as previously described.(Li et al., 2020; Millar et al., 2024; Parkinson et al., 2020) Briefly, lists were considered ranked if metrics of statistical significance (p-value) and/or fold change (FC) were reported. These lists were ordered by p-value/false discovery rate (low to high) then, where applicable, by absolute fold change or effect size (high to low). Gene names were converted to HGNC gene symbols (or Ensembl/Refseq symbols if no HGNC symbol). Rodent data was mapped to human ortholog symbols using custom scripts to allow for comparison with the human MAIC results.

***MAIC***

The MAIC algorithm has previously been described in detail.(Li et al., 2020; Millar et al., 2024; Parkinson et al., 2020; The GenOMICC Investigators et al., 2021; B. Wang et al., 2022) A full description and the source code is available at https://baillielab.net/maic. We implemented pymaic v0.2 in Python v3.9 and used Technique to categorise input lists. MAIC combines both ranked and unranked lists, of unknown quality, to build a comprehensive ranked list of entities according to 4 basic assumptions.1. There is a set of true positives (genes implicated in BPD), 2. A gene is more likely to be a true positive if it appears in datasets from more than one source, 3. A gene is more likely to be a true positive if it appears in datasets with a higher proportion of replicated genes. 4. A gene is more likely to be a true positive if it appears in datasets from multiple methods or modalities.

***Functional analyses***

All enrichment and over representation analyses were implemented using clusterProfiler (v4.0)(Wu et al., 2021) in R (v4.4.0) and visualised using functions in that package. Redundancy of enriched GO terms was removed using the simplify function. Hypergeometric tests were implemented with the 1-phyperfunction in R. Gene overlaps were visualised using ggvenn.

***Protein Interaction Network***

Prioritised genes were analysed using the online tool STRING (https:// string-db.org) to determine potential protein-protein interactions. For humans 919 genes were mapped, for rodents 1726. A PPI network was constructed using the MCL (Markov Clustering) algorithm, with an inflation parameter of 3 and focused on high confidence (interaction score of >= 0.7) interactions . The network was exported and hub genes subsequently identified using CytoHubba 0.1 implemented in CytoScape 3.10.3. The overlap of the top 100 ranked genes, ranked using 5 common algorithms (MCC, MNC, Degree, EPC and DMNC), was used to evaluate hub genes, apart from overlaps where the top 10 were used.

**References**

Ahmed, S., Odumade, O. A., van Zalm, P., Fatou, B., Hansen, R., Martin, C. R., Angelidou, A., & Steen, H. (2023). Proteomics-Based Mapping of Bronchopulmonary Dysplasia-Associated Changes in Noninvasively Accessible Oral Secretions. *The Journal of Pediatrics*, *270*. Ovid MEDLINE(R) Epub Ahead of Print. https://doi.org/10.1016/j.jpeds.2023.113774

Ambalavanan, N., Cotten, C. M., Page, G. P., Carlo, W. A., Murray, J. C., Bhattacharya, S., Mariani, T. J., Cuna, A. C., Faye-Petersen, O. M., Kelly, D., & Higgins, R. D. (2015). Integrated Genomic Analyses in Bronchopulmonary Dysplasia. *The Journal of Pediatrics*, *166*(3), 531-537.e13. https://doi.org/10.1016/j.jpeds.2014.09.052

Bhandari, V., Bizzarro, M. J., Shetty, A., Zhong, X., Page, G. P., Zhang, H., Ment, L. R., Gruen, J. R., & for the Neonatal Genetics Study Group. (2006). Familial and Genetic Susceptibility to Major Neonatal Morbidities in Preterm Twins. *Pediatrics*, *117*(6), 1901–1906. https://doi.org/10.1542/peds.2005-1414

Bhattacharya, S., Go, D., Krenitsky, D. L., Huyck, H. L., Solleti, S. K., Lunger, V. A., Metlay, L., Srisuma, S., Wert, S. E., Mariani, T. J., & Pryhuber, G. S. (2012). Genome-wide transcriptional profiling reveals connective tissue mast cell accumulation in bronchopulmonary dysplasia. *American Journal of Respiratory and Critical Care Medicine*, *186*(4). Ovid MEDLINE(R) <2012>. https://doi.org/10.1164/rccm.201203-0406OC

Bhattacharya, S., Mereness, J. A., Baran, A. M., Misra, R. S., Peterson, D. R., Ryan, R. M., Reynolds, A. M., Pryhuber, G. S., & Mariani, T. J. (2020). Lymphocyte-Specific Biomarkers Associated With Preterm Birth and Bronchopulmonary Dysplasia. *Frontiers in Immunology*, *11*. Ovid MEDLINE(R) <2020>. https://doi.org/10.3389/fimmu.2020.563473

Clark, J., Glasziou, P., Del Mar, C., Bannach-Brown, A., Stehlik, P., & Scott, A. M. (2020). A full systematic review was completed in 2 weeks using automation tools: A case study. *Journal of Clinical Epidemiology*, *121*, 81–90. https://doi.org/10.1016/j.jclinepi.2020.01.008

Dassios, T., & Greenough, A. (2021). Long-term sequelae of bronchopulmonary dysplasia. In I. P. Sinha, J. M. Bhatt, A. Cleator, & H. Wallace (Eds), *Respiratory Diseases of the Newborn Infant* (pp. 68–78). European Respiratory Society. https://doi.org/10.1183/2312508X.10013720

Eiby, Y. A., Wright, L. L., Kalanjati, V. P., Miller, S. M., Bjorkman, S. T., Keates, H. L., Lumbers, E. R., Colditz, P. B., & Lingwood, B. E. (2013). A Pig Model of the Preterm Neonate: Anthropometric and Physiological Characteristics. *PLoS ONE*, *8*(7), e68763. https://doi.org/10.1371/journal.pone.0068763

Gilfillan, M., Bhandari, A., & Bhandari, V. (2021). Diagnosis and management of bronchopulmonary dysplasia. *BMJ*, n1974. https://doi.org/10.1136/bmj.n1974

Greenough, A. (2000). Measuring respiratory outcome. *Seminars in Neonatology*, *5*(2), 119–126. https://doi.org/10.1053/siny.1999.0006

Greenough, A. (2006). Bronchopulmonary dysplasia – Long term follow up. *Paediatric Respiratory Reviews*, *7*, S189–S191. https://doi.org/10.1016/j.prrv.2006.04.206

Hadchouel, A., Durrmeyer, X., Bouzigon, E., Incitti, R., Huusko, J., Jarreau, P.-H., Lenclen, R., Demenais, F., Franco-Montoya, M.-L., Layouni, I., Patkai, J., Bourbon, J., Hallman, M., Danan, C., & Delacourt, C. (2011). Identification of *SPOCK2* As a Susceptibility Gene for Bronchopulmonary Dysplasia. *American Journal of Respiratory and Critical Care Medicine*, *184*(10), 1164–1170. https://doi.org/10.1164/rccm.201103-0548OC

Hurskainen, M., Mizikova, I., Cook, D. P., Andersson, N., Cyr-Depauw, C., Lesage, F., Helle, E., Renesme, L., Jankov, R. P., Heikinheimo, M., Vanderhyden, B. C., & Thebaud, B. (2021). Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. *Nature Communications*, *12*(1). Ovid MEDLINE(R) <2021>. https://doi.org/10.1038/s41467-021-21865-2

Isayama, T., Lee, S. K., Yang, J., Lee, D., Daspal, S., Dunn, M., Shah, P. S., & for the Canadian Neonatal Network and Canadian Neonatal Follow-Up Network Investigators. (2017). Revisiting the Definition of Bronchopulmonary Dysplasia: Effect of Changing Panoply of Respiratory Support for Preterm Neonates. *JAMA Pediatrics*, *171*(3), 271. https://doi.org/10.1001/jamapediatrics.2016.4141

Jensen, E. A., Dysart, K., Gantz, M. G., McDonald, S., Bamat, N. A., Keszler, M., Kirpalani, H., Laughon, M. M., Poindexter, B. B., Duncan, A. F., Yoder, B. A., Eichenwald, E. C., & DeMauro, S. B. (2019). The Diagnosis of Bronchopulmonary Dysplasia in Very Preterm Infants. An Evidence-based Approach. *American Journal of Respiratory and Critical Care Medicine*, *200*(6), 751–759. https://doi.org/10.1164/rccm.201812-2348OC

Johnson, D. R., O’Connor, J. C., Hartman, M. E., Tapping, R. I., & Freund, G. G. (2007). Acute hypoxia activates the neuroimmune system, which diabetes exacerbates. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(5), 1161–1166. https://doi.org/10.1523/JNEUROSCI.4560-06.2007

Lavoie, P. M., Pham, C., & Jang, K. L. (2008). Heritability of Bronchopulmonary Dysplasia, Defined According to the Consensus Statement of the National Institutes of Health. *Pediatrics*, *122*(3), 479–485. https://doi.org/10.1542/peds.2007-2313

Li, B., Clohisey, S. M., Chia, B. S., Wang, B., Cui, A., Eisenhaure, T., Schweitzer, L. D., Hoover, P., Parkinson, N. J., Nachshon, A., Smith, N., Regan, T., Farr, D., Gutmann, M. U., Bukhari, S. I., Law, A., Sangesland, M., Gat-Viks, I., Digard, P., … Hacohen, N. (2020). Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. *Nature Communications*, *11*(1), 164. https://doi.org/10.1038/s41467-019-13965-x

Mahlman, M., Karjalainen, M. K., Huusko, J. M., Andersson, S., Kari, M. A., Tammela, O. K. T., Sankilampi, U., Lehtonen, L., Marttila, R. H., Bassler, D., Poets, C. F., Lacaze-Masmonteil, T., Danan, C., Delacourt, C., Palotie, A., Muglia, L. J., Lavoie, P. M., Hadchouel, A., Rämet, M., & Hallman, M. (2017). Genome-wide association study of bronchopulmonary dysplasia: A potential role for variants near the CRP gene. *Scientific Reports*, *7*(1), 9271. https://doi.org/10.1038/s41598-017-08977-w

Millar, J. E., Clohisey-Hendry, S., McMannus, M., Zechner, M., Wang, B., Parkinson, N., Jungnickel, M., Zaki, N. M., Pairo-Castineira, E., Rawlik, K., Rogers, J., Russell, C. D., Bos, L. D., Meyer, N. J., Calfee, C., McAuley, D. F., Shankar-Hari, M., & Baillie, J. K. (2024). *The genomic landscape of Acute Respiratory Distress Syndrome: A meta-analysis by information content of genome-wide studies of the host response*. https://doi.org/10.1101/2024.02.13.24301089

Moreira, A. G., Arora, T., Arya, S., Winter, C., Valadie, C. T., & Kwinta, P. (2023). Leveraging transcriptomics to develop bronchopulmonary dysplasia endotypes: A concept paper. *Respiratory Research*, *24*(1). Ovid MEDLINE(R) <2023 to 2024>. https://doi.org/10.1186/s12931-023-02596-y

*Neonatal Data Analysis Unit (NDAU). Neonatal Health Intelligence Tool. 2021. Available: Https://www.imperial.ac.uk/neonatal-data-analysis-unit/neonatal-data-analysis-unit/neonatal-data-visualisations/*. (n.d.).

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Parkinson, N., Rodgers, N., Head Fourman, M., Wang, B., Zechner, M., Swets, M. C., Millar, J. E., Law, A., Russell, C. D., Baillie, J. K., & Clohisey, S. (2020). Dynamic data-driven meta-analysis for prioritisation of host genes implicated in COVID-19. *Scientific Reports*, *10*(1), 22303. https://doi.org/10.1038/s41598-020-79033-3

Rozance, P. J., Seedorf, G. J., Brown, A., Roe, G., O’Meara, M. C., Gien, J., Tang, J.-R., & Abman, S. H. (2011). Intrauterine growth restriction decreases pulmonary alveolar and vessel growth and causes pulmonary artery endothelial cell dysfunction in vitro in fetal sheep. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, *301*(6), L860–L871. https://doi.org/10.1152/ajplung.00197.2011

Sun, T., Yu, H.-Y., Yang, M., Song, Y.-F., & Fu, J.-H. (2023). Risk of asthma in preterm infants with bronchopulmonary dysplasia: A systematic review and meta-analysis. *World Journal of Pediatrics*, *19*(6), 549–556. https://doi.org/10.1007/s12519-023-00701-1

The GenOMICC Investigators, The ISARIC4C Investigators, The COVID-19 Human Genetics Initiative, 23andMe Investigators, BRACOVID Investigators, Gen-COVID Investigators, Pairo-Castineira, E., Clohisey, S., Klaric, L., Bretherick, A. D., Rawlik, K., Pasko, D., Walker, S., Parkinson, N., Fourman, M. H., Russell, C. D., Furniss, J., Richmond, A., Gountouna, E., … Baillie, J. K. (2021). Genetic mechanisms of critical illness in COVID-19. *Nature*, *591*(7848), 92–98. https://doi.org/10.1038/s41586-020-03065-y

Thébaud, B., Goss, K. N., Laughon, M., Whitsett, J. A., Abman, S. H., Steinhorn, R. H., Aschner, J. L., Davis, P. G., McGrath-Morrow, S. A., Soll, R. F., & Jobe, A. H. (2019). Bronchopulmonary dysplasia. *Nature Reviews Disease Primers*, *5*(1), 78. https://doi.org/10.1038/s41572-019-0127-7

Torgerson, D. G., Ballard, P. L., Keller, R. L., Oh, S. S., Huntsman, S., Hu, D., Eng, C., Burchard, E. G., Ballard, R. A., & TOLSURF Study Group. (2018). Ancestry and genetic associations with bronchopulmonary dysplasia in preterm infants. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, *315*(5), L858–L869. https://doi.org/10.1152/ajplung.00073.2018

Wang, B., Law, A., Regan, T., Parkinson, N., Cole, J., Russell, C. D., Dockrell, D. H., Gutmann, M. U., & Baillie, J. K. (2022). Systematic comparison of ranking aggregation methods for gene lists in experimental results. *Bioinformatics*, *38*(21), 4927–4933. https://doi.org/10.1093/bioinformatics/btac621

Wang, H., St. Julien, K. R., Stevenson, D. K., Hoffmann, T. J., Witte, J. S., Lazzeroni, L. C., Krasnow, M. A., Quaintance, C. C., Oehlert, J. W., Jelliffe-Pawlowski, L. L., Gould, J. B., Shaw, G. M., & O’Brodovich, H. M. (2013). A Genome-Wide Association Study (GWAS) for Bronchopulmonary Dysplasia. *Pediatrics*, *132*(2), 290–297. https://doi.org/10.1542/peds.2013-0533

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, *2*(3), 100141. https://doi.org/10.1016/j.xinn.2021.100141

**CONFLICT OF INTEREST**

All authors report no conflicts of interest.

**FUNDING**

<INSERT>

**DATA AVAILABILITY**

The MAIC output is included as supplementary material. All code is available at https://github.com/baillielab/bpd\_maic.

**CONTRIBUTIONS**

SCH, JKB and JEM conceived the study. SCH, PK, CH, JAR, NM, AA, NP, EL, MP manually reviewed abstracts for inclusion. SCH, PK, CH and CS curated the data. SCH, PK and CH did the formal analysis. SCH supervised the study. SCH, PK and CH wrote the original draft of the manuscript. All authors reviewed and edited the manuscript. SCH validated the study data. SCH, PK and CH had access to the raw data. The corresponding author had full access to all the data and final responsibility for the decision to submit for publication.