NIHR | National Institute for Health and Care Research

PROSPERO
International prospective register of systematic reviews

Prioritising host genes implicated in Bronchopulmonary Dysplasia: a systematic review and meta-analysis by information content of omic studies

Sara Clohisey Hendry, J. Kenneth Baillie, Jonathan Millar, Prerna Khanna

Citation
Sara Clohisey Hendry, J. Kenneth Baillie, Jonathan Millar, Prerna Khanna. Prioritising host genes implicated in Bronchopulmonary Dysplasia: a systematic review and meta-analysis by information content of omic studies. PROSPERO 2024 Available from https://www.crd.york.ac.uk/PROSPERO/view/CRD42024550229

REVIEW TITLE AND BASIC DETAILS

Review title
Prioritising host genes implicated in Bronchopulmonary Dysplasia: a systematic review and meta-analysis by information content of omic studies

Review objectives
In premature babies with Bronchopulmonary Dysplasia, which genes are prioritised as being implicated in the host-response?

Keywords
Bronchopulmonary Dysplasia, Genetics, Genomics, Hyaline Membrane Disease, Respiratory Distress Syndrome, Susceptibility

SEARCHING AND SCREENING

Searches
Primary search
OVID MEDLINE and Epub Ahead of Print, In-Process, In-Data Review & Other Non-Indexed Citations, Daily and Versions 1946 to January 31, 2024.
AND
Embase Classic + Embase 1947 to January 31, 2024.
Secondary search
- Search of [bioRxiv](https://www.biorxiv.org) and [medRiv](https://www.medrxiv.org) using relevant keywords.

- Single-level forward and backward citation searches of the reference lists of included articles.
- Extraction of references contained in the [ARDS DB] (https://doi.org/10.3389/fgene.2021.750568).
- Correspondence with corresponding authors of included studies.
No language restrictions will be imposed.

Study design
We are seeking to include whole-genome studies that report an association between genes, transcripts, or proteins and susceptibility to BPD or with severity or outcome.
We will include studies employing the following methodologies: CRISPR screen, RNAi screen, protein-protein interaction, host proteins incorporated into virion or virus-like particle, genome-wide association, transcriptomic study, or proteomic study.
We will exclude studies in which only results from non-human cells or animals are reported.
We will exclude studies employing the following methodologies: candidate transcriptomic or proteomic studies (< 50 genes investigated) and candidate gene association studies.
We will exclude studies including fewer than 5 patients in any arm.

ELIGIBILITY CRITERIA

Condition or domain being studied
Bronchopulmonary Dysplasia.

Population
Inclusion:
- Human studies: *in-vivo* or *in-vitro*
- Premature humans (<32 weeks PMA)
- Accepted methodologies:
- CRISPR screen
- RNAi screen
- Protein-protein interaction study
- Host proteins incorporated into virion or virus-like particle
- Genome-wide association study
- Transcriptomic study
- Proteomic study
Exclusion:
- Term babies, Children, Adults
- Animal studies
- Meta-analyses, *in-silico* analyses, or re-analysis of previously published data
- Excluded methodologies:
- Candidate *in-vivo* or *in-vitro* transcriptomic or proteomic studies (defined as those investigating < 50 genes)
- Candidate gene association studies
- Studies including fewer than 5 individuals in either the control or BPD arm

Intervention(s) or exposure(s)
Not applicable.

Comparator(s) or control(s)

Not applicable.

Context
Where applicable we will report our review in concordance with the [Human Genome Epidemiology Network (HuGE Net) Handbook of Systematic Reviews](https://www.cdc.gov/genomics/hugenet/participate.htm).

OUTCOMES TO BE ANALYSED

Main outcomes
Ranked list of genes associated with BPD susceptibility, severity, and/or outcome.

Measures of effect
Not applicable.

Additional outcomes
None

DATA COLLECTION PROCESS

Data extraction (selection and coding)
Selection of studies
Article titles and abstracts obtained using the search strategy will be stored using reference management software (endNote X9, Clarivate Analytics, United States). Intital screening of titles wil be conducted by single authors against eligibility criteria, using the Screenatron tool (Systematic Review Accelerator, Bond University, Australia). Thereafter, screening of abstracts against eligibility criteria will be conducted by two authors independently. Inconsistencies wll be resolved in discussion with a thrd author. Full text articles will be retieved for studies matching the eligibility criteria.
Data extraction
Data will be extracted by two independent reviewers using a pre-piloted proforma.
Gene lists will be extracted and their ranking preserved if possible. Ranking may be based on magnitude of effect or signficance. Where multiple measures are available we will preference magnitude of effect. Similarly, adjusted P values will be preferred over raw P values. If studies report multiple time points we will rank genes based on their minimum P value. We will exlcude genes for which the magnitude of effect or significance fall outside the authors threshold, or when this information is not available, for which $P > 0.05$, or z score < 1.96, or log fold change < 1.5. Gene, transcript, or protein identifiers will be mapped to its HUGO Gene Nomenclature Committee (HGNC) symbol. If one is not available we will use an equivalent Ensembl or Refseq symbol.
In addition, we will extract information relating to study design, methodology, tissue/cell type, demographics, BPD aetiology, risk factors, severity, and outcomes.

Risk of bias (quality) assessment
All genome-wide association studies will be assessed for risk of bias using domain-based evaluation as described in the [Q-Genie tool](https://doi.org/10.1186/s12863-015-0211-2). Studies will be classified as low, moderate, or high quality from their overall score.
For each gene ranked in the "top 50" by our meta-analysis we will rate the cumulatative

evidence for genetic association using the [Venice interim guidelines] (https://doi.org/10.1093/ije/dym159).

PLANNED DATA SYNTHESIS

Strategy for data synthesis

MAIC

We will conduct a meta-analysis by information content (MAIC) of extracted gene lists. We have previously described our MAIC methodolgy in detail (, ,). All components of our core algorithm can be found at .

Functional enrichment analysis

We will conduct gene set enrichment analysis based on rankings by MAIC score. We will use two methods: 1. FGSEA in R () using the full ranked list, and 2. over-enrichment analysis using Enrichr () on the "top 100" genes.

We will control for false discovery using the Benjamini-Hochberg procedure (FDR < 0.05).

Analysis of subgroups or subsets

If sufficient data are available we will conduct sub-group analyses based on BPD aetiology e.g., PMA at birth, viral infection, severity, death

If sufficient data are available we will conduct sub-group analyses based on ancestry.

REVIEW AFFILIATION, FUNDING AND PEER REVIEW

Review team members

- Dr Sara Clohisey Hendry, The University of Edinburgh
- Professor J. Kenneth Baillie, The University of Edinburgh
- Dr Jonathan Millar, The University of Edinburgh
- Miss Prerna Khanna, The University of Edinburgh

Review affiliation

The University of Edinburgh

Funding source

None

Named contact

Sara Clohisey Hendry. The Roslin Institute, Easter Bush Campus, EH25 9RG, Edinburgh, Scotland.

sara.clohisey@roslin.ed.ac.uk

TIMELINE OF THE REVIEW

Review timeline

Start date: 01 June 2024. End date: 31 May 2025

Date of first submission to PROSPERO

23 May 2024

Date of registration in PROSPERO
29 May 2024

CURRENT REVIEW STAGE

Publication of review results
The intention is to publish the review once completed.The review will be published in English

Stage of the review at this submission

Review stage	Started	Completed
Pilot work		
Formal searching/study identification		
Screening search results against inclusion criteria		
Data extraction or receipt of IP		
Risk of bias/quality assessment		
Data synthesis		

Review status
The review is currently planned or ongoing.

ADDITIONAL INFORMATION

PROSPERO version history

- Version 1.0 published on 29 May 2024

Review conflict of interest
None known

Country
Scotland

Medical Subject Headings
Bronchopulmonary Dysplasia; Gestational Age; Humans; Infant; Infant, Newborn; Infant, Premature

Details of any existing review of the same topic by the same authors
Our group has previously employed similar methodology to study the host genetics of [Influenza A](https://doi.org/10.1038/s41467-019-13965-x), [SARS-CoV-2](https://doi.org/10.1038/s41598-020-79033-3) infection, and [ARDS](https://www.medrxiv.org/content/10.1101/2024.02.13.24301089v1).

Disclaimer
The content of this record displays the information provided by the review team. PROSPERO does not peer review registration records or endorse their content.

PROSPERO accepts and posts the information provided in good faith; responsibility for record content rests with the review team. The owner of this record has affirmed that the information

provided is truthful and that they understand that deliberate provision of inaccurate information may be construed as scientific misconduct.

PROSPERO does not accept any liability for the content provided in this record or for its use. Readers use the information provided in this record at their own risk.

Any enquiries about the record should be referred to the named review contact