S1Q2 - Literature Search Notes

> Now find real contact matrix data and stratify the basic SIR model with the appropriate number of age groups to match the data found.

First attempt - Brian Bockelman:

- Found a paper about contact matrixes in SIR models (not in COVID dataset) via Google search ("contact matrix sir model"): https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4002176/
    - This paper is about age-stratified mathematical epidemic models, not COVID.
- Ran an xDD query for all articles referencing this paper: https://xdd.wisc.edu/api/articles?term=10.3934/mbe.2013.10.1475&include_highlights=true
- Top hit was "Age- and Sex-Specific Modelling of the COVID-19 Epidemic"
    - Link: https://www.medrxiv.org/content/10.1101/2020.10.06.20207951v1.full.pdf
    - Figure 4 of the paper contains contact rates for Germany in figure form.
    - Figure 4 references "Efficient Estimation of Age-Specific Social Contact Rates Between Men and Women," doi:10.1214/16-AOAS1006
        - That paper includes the matrix values as Figure 1 but it's unclear that's what we need?
    - Difficult to pull out data from heat maps so set this aside.
- Another hit was "Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK"

First attempt - Shiv:

Ran xDD query
https://xdd.wisc.edu/api/articles?dataset=xdd-covid-19&match=true&max=10&term=SIR age&additional_fields=title and got the following results

1	63cac7d174bed2df5caf87e7	10.1038/s41598-021-94609-3	A modified age-structured SIR model for COVID-19 type viruses

2	5f590c70a58f1dfd521481cc	10.1101/2020.09.07.20184887	Age-structured SIR model and resource growth dynamics: A preliminary COVID-19 study
3	616e8ddf67467f7269d455f5	10.1186/s13362-020-00090-4	An age and space structured SIR model describing the Covid-19 pandemic
4	5ec5647d998e17af826f549e	10.1101/2020.05.15.20103317	An Age and Space Structured SIR Model Describing the Covid-19 Pandemic

First result, “A modified age-structured SIR model for COVID-19 type viruses”.

- Figure 1 from 10.1038/s41598-021-94609-3 provides a contact matrix.
    - Converted this using Microsoft Excel and uploaded to Github.
- Reference

Second try

- running query “SIR contact age” with inclusive=true
- Figure 2 from https://www.medrxiv.org/content/10.1101/2020.07.05.20146647v1.full.pdf
    - Again, data is in a heat map which would be difficult to parse.
    - Data comes from reference 20, 10.1038/s41598-021-94609-3
    - Pulled up PDF from reference 20, (“Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK”), which is in xDD as https://xdd.wisc.edu/api/articles?docid=63d7ef4a74bed2df5c597cc8
    - This reference Includes a GitHub repo: https://github.com/jarvisc1/comix_covid-19-first_wave with raw data
        - Data is from a survey in the UK taken during lockdowns – may not be so useful.
        - Data converted to CSV and uploaded to GitHub.
- Another result is for 10.1016/j.eclinm.2020.100354 https://www.sciencedirect.com/science/article/pii/S2589537020300985 and Figure 2 looks very relevant

Finally, Brian returned to the “first result” of Shiv’s work.

- Figure 1 (which we liked) comes from data in reference 15, “Prem, K. & Cook, A. R. Projecting social contact matrices in 152 countries using contact surveys and demographic data. PLoS Comput. Biol. 13, 20 (2017).” (https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005697)
- The reference included additional supporting information:
    - https://doi.org/10.1371/journal.pcbi.1005697.s001 PDF explaining the data.
    - https://doi.org/10.1371/journal.pcbi.1005697.s002 Zip file of multiple Excel spreadsheets
- The Excel spreadsheets include, for ~150 countries,
    - Result of contact surveys. For survey participants, the average number of individuals in each age bin (each age bin is 5 years, up to X=16) reported having contact with.
    - The surveys additionally broke down data by contact location (home, school, work, other).
- The supporting PDF suggested reweighting the different location datasets to simulate interventions (e.g., reweight to 0 for school contacts to represent school closer, reweight to 0.5 the work and other interactions to represent social distancing).
    - We stopped here. Note that potentially the weights could be played with for the UK dataset to see how close one can get to the UK during-pandemic surveys.

- Note: We also checked in UK data to help with reweighting
-

> To simulate the model with realistic initial values, find data on population distribution by age group