SUPPORTING INFORMATION

Naturally-acquired protection against upper respiratory symptoms involving group A *Streptococcus* in a longitudinal cohort study

Joseph A. Lewnard, PhD[1,2,3,*], Lilith K. Whittles, PhD[4,5,6], Anne-Marie Rick, MD, MPH [7,8], Judith M. Martin MD[7,8]

1. Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, California
2. Division of Infectious Diseases & Vaccinology, School of Public Health, University of California, Berkeley, Berkeley, California
3. Center for Computational Biology, College of Engineering, University of California, Berkeley, Berkeley, California
4. Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK
5. MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK
6. NIHR Health Protection Research Unit in Modelling Methodology, School of Public Health, Imperial College London, London, UK
7. Department of Pediatrics, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania
8. Department of Pediatrics, UPMC Children's Hospital of Pittsburgh, University of Pittsburgh, Pittsburgh, Pennsylvania

*   Corresponding author:
     Joseph Lewnard, PhD
     2121 Berkeley Way, Office 5410
     Berkeley, CA 94720
     jlewnard@berkeley.edu
     510-664-4050

Contents

Table S1: Summary of GAS carriage and symptoms among all children

Cohort (entry year)	Symptoms at time of GAS detection					
	Typical symptoms		Atypical symptoms		Asymptomatic detections	
	Number of visits	*Number of children*	*Number of visits*	*Number of children*	*Number of visits*	*Number of children*
1998-99						
	0	20	0	30	0	18
	1	10	1	10	1-4	16
	2	5	2	5	5-9	5
	3	6	3	1	10-14	3
	4	3	4	2	15-20	1
	5	3	5	0	20-24	2
	6	1	6	0	25-29	3
1999-2000						
	0	9	0	17	0	9
	1	7	1	9	1-4	15
	2	6	2	0	5-9	3
	3	3	3	4	10-14	0
	4	2	4	0	15-20	1
	5	3	5	0	20-24	0
	6	0	6	0	25-29	1
2000-01						
	0	9	0	21	0	20
	1	13	1	9	1-4	4
	2	4	2	2	5-9	3
	3	4	3	0	10-14	3
	4	1	4	0	15-20	2
	5	1	5	0	20-24	0
2001-02						
	0	9	0	6	0	7
	1	3	1	7	1-4	5
	2	2	2	2	5-9	2
	3	1	3	0	10-14	1
	4	0	4	0	15-20	0
2002-03						
	0	16	0	17	0	13
	1	3	1	3	1-4	5
	2	1	2	0	5-9	2

Table S2: Summary of distinct GAS type detections among all children

Cohort (entry year)	Unique types detected					
	FIGE types		emm types		emm clusters	
	Number of unique types	Number of children	Number of unique types	Number of children	Number of unique clusters	Number of children
1998-99						
	0	11	0	11	0	11
	1	6	1	6	1	6
	2	9	2	10	2	15
	3	10	3	10	3	6
	4	2	4	5	4	7
	5	3	5	6	5	3
	6	5	6	0	6	0
	7	2	7	0	7	0
1999-2000						
	0	4	0	4	0	4
	1	5	1	5	1	6
	2	8	2	9	2	9
	3	5	3	7	3	7
	4	3	4	4	4	4
	5	3	5	1	5	0
	6	2	6	0	6	0
2000-01						
	0	8	0	8	0	8
	1	9	1	10	1	10
	2	6	2	11	2	11
	3	6	3	2	3	3
	4	2	4	0	4	0
	5	1	5	1	5	0
2001-02						
	0	4	0	4	0	4
	1	6	1	6	1	6
	2	2	2	3	2	3
	3	1	3	2	3	2
	4	2	4	0	4	0
2002-03						
	0	9	0	9	0	9
	1	9	1	9	1	9
	2	2	2	2	2	2

Table S3: Summary of distinct GAS type detections at visits with and without symptoms

Cohort (entry year)	Symptoms at time of GAS detection					
	Typical symptoms		Atypical symptoms		Asymptomatic detections	
	Number of FIGE types	*Number of children*	*Number of FIGE types*	*Number of children*	*Number of FIGE types*	*Number of children*
1998-99						
	0	20	0	31	0	18
	1	10	1	11	1	12
	2	7	2	4	2	6
	3	8	3	1	3	2
	4	3	4	1	4	0
	5	0	5	0	5	7
	6	0	6	0	6	2
	7	0	7	0	7	1
1999-2000						
	0	9	0	17	0	9
	1	9	1	10	1	9
	2	7	2	3	2	7
	3	4	3	0	3	1
	4	1	4	0	4	2
	5	0	5	0	5	1
	6	0	6	0	6	1
2000-01						
	0	10	0	21	0	20
	1	16	1	11	1	3
	2	4	2	0	2	3
	3	2	3	0	3	4
	4	0	4	0	4	1
	5	0	5	0	5	1
2001-02						
	0	9	0	6	0	7
	1	5	1	9	1	6
	2	1	2	0	2	1
	3	0	3	0	3	1
2002-03						
	0	16	0	17	0	13
	1	4	1	3	1	6
	2	0	2	0	2	1
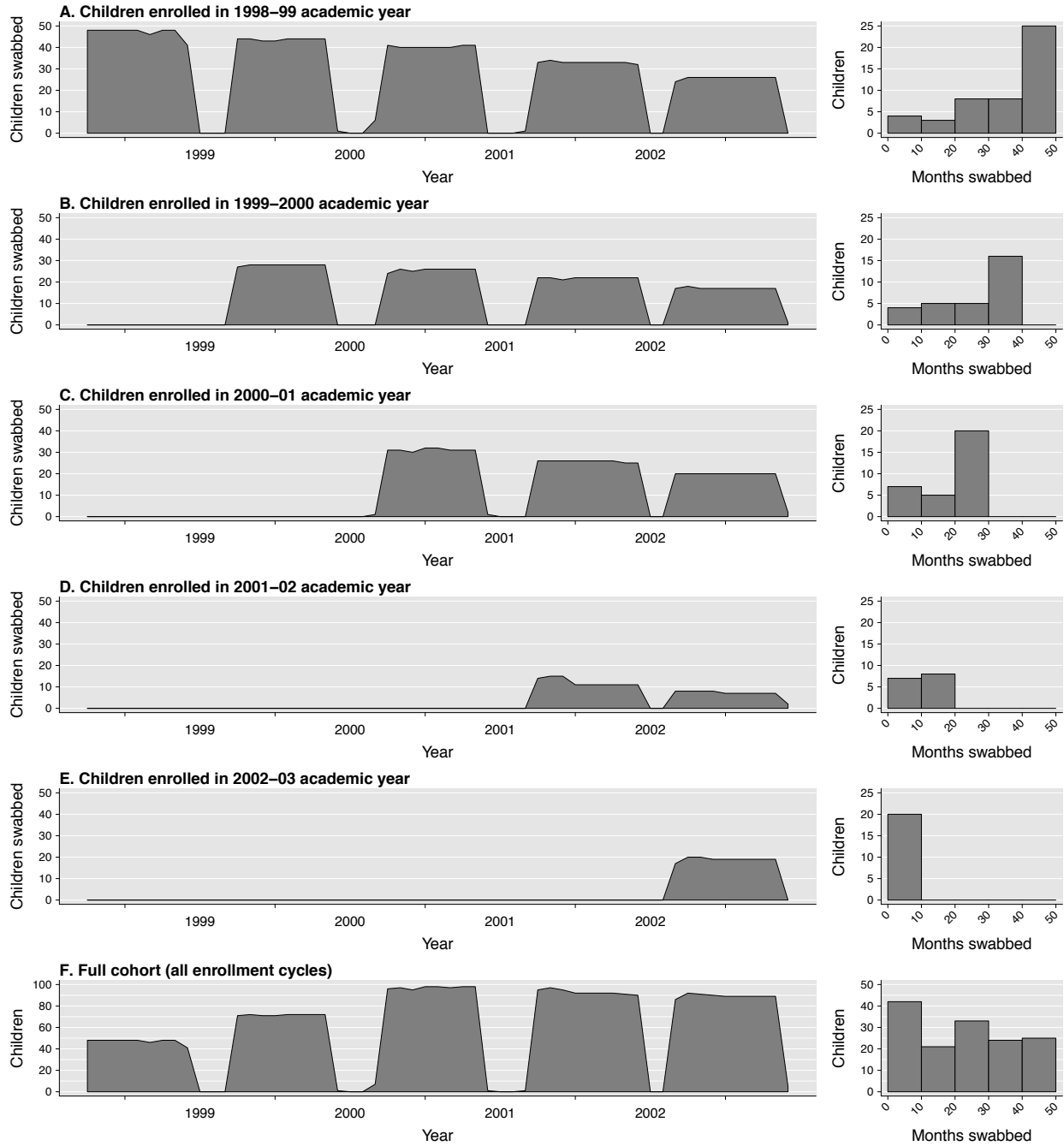
Figure S1. Cohort enrollment over time. We plot the number of children participating in the study each month (defined as having cultures obtained at least once; left panels) by year of entry (**A**–**E**), and for the full cohort (**F**), together with the distributions of follow-up durations among children (defined as the number of unique months in which they had cultures obtained at least once); right panels.
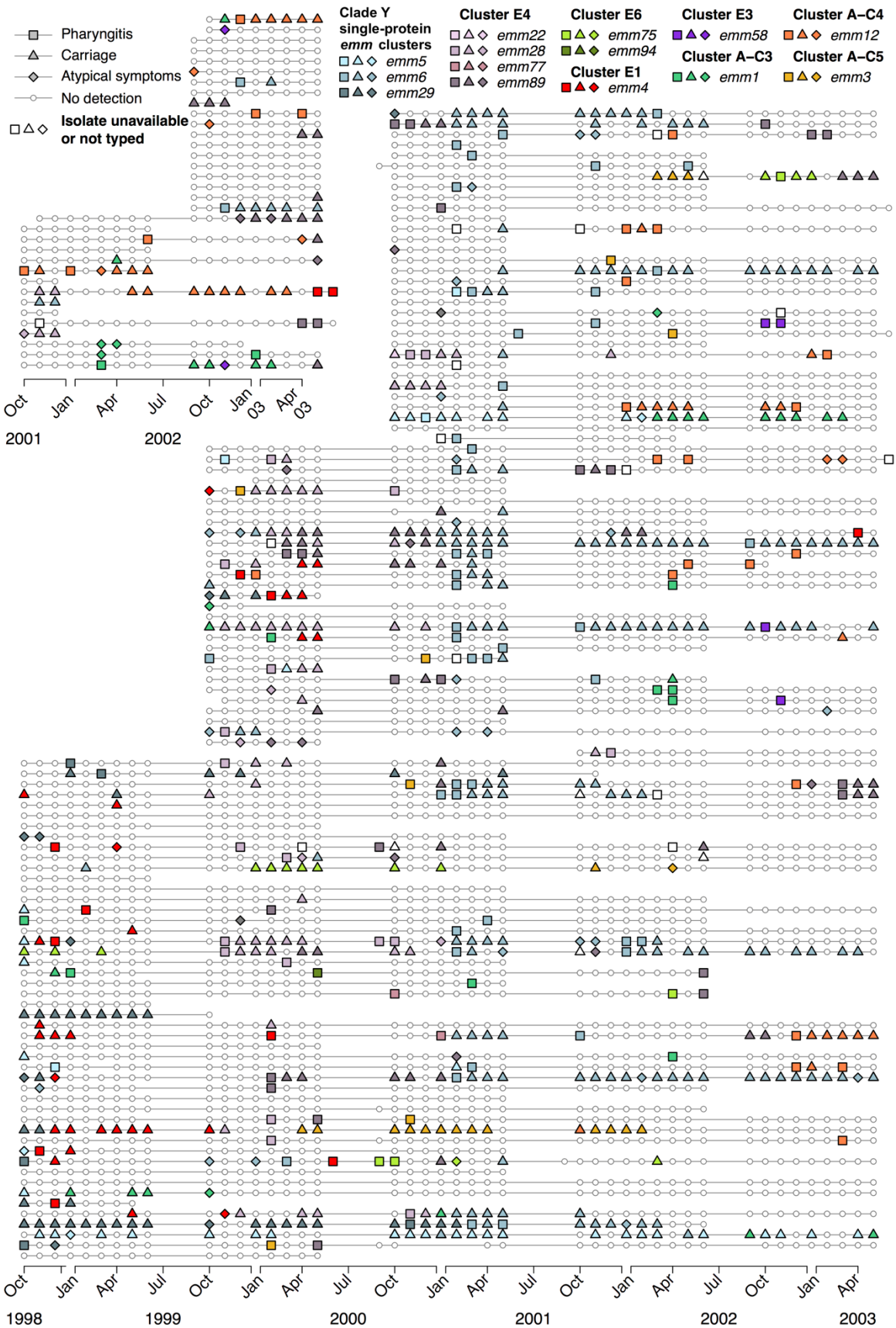
Figure S2. Simplified histories of *emm* type detections and symptoms. Each horizontal line represents a simplified time series of GAS detections, for each child, over the period of the study. For clarity, observations are aggregated by month, to illustrate months with any GAS detection (but no symptoms), months with any GAS detection with atypical symptoms, and months with any GAS detection with typical symptoms. Observations from children entering during the years 2001 and later are presented in the top-left corner to fit all data in the space of a single image. We present the total number of isolates belonging to each *emm* cluster, *emm* type, and unique FIGE type in **Table 2** of the main text. Within clade Y, *emm* types 5, 6, and 29 each represent a unique *emm* cluster comprising of a single *emm* type.