BioSocieties (2009), **4**, 207–222 © London School of Economics and Political Science
doi:10.1017/S1745855209990111

The Lives of 'Facts' in Mathematical Models: A Story of Population-level Disease Transmission of *Haemophilus Influenzae* Type B Bacteria

Erika Mansnerus

Centre for Analysis of Risk and Regulation (CARR), London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK
E-mail: e.mansnerus@lse.ac.uk

Abstract

This article studies how our understanding of population-level disease transmission has evolved over time. The main question is: what happens to 'facts' in the course of their life history? The concept of life history *captures the process that shapes the facts of disease transmission, mobilizes them via mathematical and graphical representations, and allows them to evolve and change over time. Hence, this concept provides continuity from knowledge production to utilization. Life history is developed through phases in the 'lives' of 'facts': birth and youth, adulthood and reproductive years, and old age. The life-history approach consists of a set of 'facts' binding together knowledge of a disease, its routes of transmission, and the susceptibility of the exposed population; it thus provides an adequate framework to explore the complex nature of population-level disease transmission. The analytical focus of this article is concerned with how these 'facts' are disseminated via model-based or mathematical representations. Just as life histories are stories full of interactions, surprises and struggles, this article highlights the underlying contingencies in the dissemination and accumulation of factual knowledge.*

Keywords Disease Transmission, Evidence, Facts, Knowledge Transfer, Models, Public Health

Introduction: a life history of 'facts'

How did our understanding of population-level disease transmission evolve over time? This article argues on the side of continuity between narratives of knowledge production and utilization by elaborating a concept of *life histories* of 'facts' of disease transmission. Simulations of disease transmission depict multiple aspects of the process: the reservoir of the

Erika Mansnerus works as an ESRC Postdoctoral Fellow at the Centre for Analysis of Risk and Regulation (CARR) at the London School of Economics and Political Science (LSE). Prior to this, she was a Postdoctoral Research Officer in a Leverhulme Trust/ESRC funded project 'Nature of Evidence: How Well Do "Facts" Travel?' Economic History Department, LSE. Her research focuses on the use of modelling techniques in the governance of public health risks.

disease, the route of infection, the mechanisms of prevention and the impact of human interaction. In describing this complex sequence, simulation models employ mathematical representations to describe the disease processes in concise analytical terms. Each step of this process introduces multiple 'facts'.[1] Examining the exchange of information between communities, this paper investigates how our understanding of disease transmission evolves over time.

These aspects will be explored by developing *a life history of facts* as a framework, which invites us to follow the different phases in the growth of knowledge: from birth and youth, through adulthood and the reproductive years to old age. In a way, the growth of knowledge expresses the continuity between the processes of knowledge production and utilization. The case study explored in this article concerns population-level transmission dynamics of *Haemophilus influenzae* type b bacteria (Hib) and the public health measures that aim to protect the public from life-threatening diseases caused by this bacteria, such as meningitis or septicaemia.

The starting point of a disease's 'life history' is the moment when epidemiological facts are discovered or *born*. It is at this moment that a specific pathogen, understood to be the singular cause of a disease, is related to mathematical representations of population-level transmission. The birth of the 'facts' will, therefore, be traced to efforts to mathematically represent population-level disease transmission.

The adulthood and reproductive years (to continue with our biographical metaphor) of a disease outbreak will then be discussed in relation to how 'facts' are *reproduced* in simulation models. The framework of the *life-history* study is based on how disease transmission is expressed in a set of detailed Hib transmission models[2] and considers the maturation of our understanding of disease transmission.

Hence, the conceptual framework of a *life history* of 'facts' will show how knowledge is not only produced, but how it organically changes and evolves over time and how model-based representations, given the complex nature of population-level transmission dynamics, enable the maturation of knowledge.

To develop the notion of a *life history* of 'facts', this study combines historical and sociological perspectives. It follows Lorraine Daston's notion of biographies of scientific objects, which she uses to create a framework for studying—*vita activa*—the coming into being and passing away of scientific objects (Daston, 2000). The metaphor of life is applied as a concept that allows us to observe changes in the objects of research (Creager, 2002; Mendelsohn, 2003). The notion of *life history* is also similar to the notion of *trajectory*, which Hans-Jörg Rheinberger (2000: 273) uses to describe how: 'scientific objects come into existence as a result of unprecedented events [...] and they remain objects of research as long as they have the power to manifest themselves in yet unthought-of ways in the

1 Throughout this article, the concept of a 'fact' is understood as a knowledge claim accepted within a community and found to be reliable by its members. This community perspective on 'facts' aims to free our thinking from the propositional status of knowledge claims and underline the usability and applicability of 'facts'. A similar account is presented by Becker, who claims that: 'facts are facts only when they are accepted as such by the people to whom those facts are relevant' (Becker, 2007: 12–13). In order to avoid a relativistic reading of 'facts', they maintain a degree of integrity when utilized by different communities. This argument is elaborated in Morgan (forthcoming) and Mansnerus (forthcoming).

2 This group of models was built in collaboration between the University of Helsinki, the National Public Health Institute and the Helsinki University of Technology between 1994 and 2003.

future'. Even though this article is not about scientific *objects* but *facts*, this framework can similarly be used to reconstruct the *vita activa* of a fact, tracing the 'unprecedented events' in which such facts are materialized, described or conceptualized; and following them as they manifest themselves in different ways, and in different contexts or domains of research or research-inspired activities. Yet, *life history* moves a step further than these two notions—*life history* brings forth the different phases, and the evolution of these phases, which are irrevocably linked to discussions concerning nature and organic life. In other words, the notion suggests continuity from knowledge production to its utilization. It allows us, then, to observe fluidity between knowledge production and utilization across time, which is useful in order to understand how a complex phenomenon of disease transmission evolves.

The chosen perspective in this article, the life history of a fact, shows us that current understandings of population-level disease transmission fluctuate between different scholarly traditions and approaches. Life history not only gives us a vivid framework to assess the different phases of knowledge acquisition and dissemination across time and within communities, it also expands the constructivist account on knowledge production, which brings to bear contingencies, surprises, changes and forgotten moments in the dissemination of scientific knowledge (see Knorr Cetina, 1981; Latour & Woolgar, 1986 [1979]).

On methods and data

The analysis employed in this study is based on multiple sources of data in order to capture the time dimension necessary for a life history. Historical studies of epidemiology are used as secondary literature to locate early developments in the understanding of disease transmission and immunity. This material is used for analysing the *birth and youth* of transmission 'facts'. Epidemiology textbooks, in particular, are useful for identifying trends in mathematical epidemiology, which introduces us to the general acceptance of modelling methods within this field. The mathematical representations of transmission, documented in these textbooks, allow us to consider the broader framework of using mathematical tools to represent public health phenomena. A detailed analysis on the maturation of facts through modelling activities conducted by a Finnish research group is based on a more complete set of data.[3] Scientific publications from the group are the primary resource for locating and identifying the facts about transmission, whereas analysis of the interactive

3 This analysis is based on scientific publications between 1994 and 2003, and interviews and ethnographic observations in a series of workshops ($n = 22$, held in 2002–2003 at the National Public Health Institute in Helsinki). The main aim of these activities was to combine documentary data with interactional observations in order to show how facts are domesticated in models and what kind of reasoning, support and argumentation this process requires. The meetings ($n = 5$) were chosen from the main body of data on the basis of field reports. The selected meetings were then analysed by looking at the topics discussed in each meeting and tracing the dialogue on transmission. Related topics, including carriage, immunity and data acquisition, were also located. After categorizing the discussions recorded at the meetings, the transcripts were read and studied carefully, the discussions were linked with other available data from the meetings (versions of models, drafts of articles, etc.) and the context of the discussions was analysed by focusing on the 'facts' presented on disease transmission and by the impressions of participants—whether a representation or definition of a detail on transmission dynamics was clear to the participants or whether it required further explanation. Also, the ways in which the transmission dynamics were 'chopped' in order to be tailored to the model were followed. This was often expressed in terms of searching and defining the optimal parametric values and validating the choice in discussion with epidemiologists. The use of multiple datasets is also a form of triangulation, a process of thickening the description with various sources of data and testing these sources against each other.

modelling practices provides insight into the ways in which the general facts in epidemiology and statistics are specified in model-based representations.

Therefore this article follows the historical development of our understanding of disease transmission—providing a selective socio-historical narrative on the ways in which singular facts regarding microbial mechanisms of disease are related to facts describing the dynamics of disease transmission in a population. Population-level disease transmission is a '*big* fact'. A big fact is, as Sabina Leonelli (forthcoming) defines it, an end-result of research. It provides a general description of a phenomenon and is formulated propositionally. Mary Morgan (forthcoming) broadens Leonelli's account by characterizing 'facts':

> Facts may indeed be relatively short and concrete, but our research gave rise to another observation: facts do come in various 'sizes'. Facts can be little: observations on the buds of a flowering plant; singular: the age of a particular person; come in crowds: infection rates; or be quite generic: the alpha-male in romance fiction or the exit pattern of firms in declining industries.

Population-level disease transmission is a big fact, since it appears as an end result of research, consisting of the various direct and indirect observations related to the dynamics of transmission. Thus, in this article, the big fact is analysed through its constituents, which form a set of three 'facts' that 'live their lives' through the development of epidemiological theory. These facts (1) identify Hib, (2) clarify how to acquire immunity against severe forms of the disease caused by the Hib bacteria and to show the indirect, population-level effects of immunity, and (3) describe the specific transmission dynamics used in probabilistic models. By following this set of facts within the analytical framework of life histories, this study documents the accumulation of knowledge and the stabilization of factual claims through the mathematical representations employed in epidemiological studies.

The structure of this article follows the phases in a life history: birth and youth, adulthood and reproductive years, and old age. The next section explores the birth and youth of disease transmission. It introduces the processes of identifying the causative agents of Hib diseases and explores how the public health perspective was incorporated into the studies. The following section analyses how facts of transmission patterns are represented and reproduced in models, and reveals how these facts utilize the models' capacity to simplify the complexity of disease patterns into a flexible form. The penultimate section discusses how facts become 'seniors'—that is, 'old age facts'—and how they act as *expert facts* in order to disseminate previously established knowledge claims in novel contexts. The final section discusses why life history is a useful framework to analyse the dissemination of factual knowledge across time and disciplinary communities.

Birth and youth of 'facts': from identification of the bacteria to its population-level effects

Building our understanding of disease transmission is a gradual process. In this section, I will explore how the facts used to identify Hib, and the facts used to clarify its transmission in a population, were discovered. I call this emergence of facts *birth and youth*, and it represents the first phase in the *life history* of facts. I will provide a comprehensive picture

of the early phases of Hib studies from a public health perspective. This perspective is concerned with the severe forms of disease caused by the pathogen in a population and is expressed as a *big fact*—a fact that comprises various aspects of a phenomenon or is built upon the 'end-results' of research. In a way, a life history of a fact shows how understanding a complex phenomenon develops gradually, through the unfolding of 'facts' that define the process. Moreover, I will show how the set of facts concerning population-level transmission of Hib diseases (F1–F3) emerge in epidemiological studies. The set of facts that are of greatest interest to this study are:

- F1: a fact that identifies *Haemophilus influenzae* and its strains (particularly strain b, i.e. Hib), and describes the correlation between the bacteria and the severe forms of disease it may cause.
- F2: a fact that clarifies immunity to Hib and shows its indirect, population-level effects.
- F3: a fact that explains the transmission dynamics of Hib within a population.

Understanding germs: identifying Hib

In order to understand how population-level transmission dynamics evolved over time and to conceptualize it as a life history of a fact, I will first introduce the discovery and identification of the pathogen, namely, *Haemophilus influenzae* type b bacteria, in the context of germ theory. The main focus of germ theory was on how transmission occurs between individuals. However, this study is primarily concerned with population-level transmission dynamics.

The formulation of germ theory created a new framework to identify factors behind infectious disease. At the core of this framework are Koch's postulates,[4] which introduced the generalized principles with which to examine whether an organism can be accepted as a cause of a particular disease. This had important implications for epidemiological study. First, it paved the way for the development of vaccines and understanding immunity. Second, it made it possible to specify the concepts that describe transmission dynamics. Understanding transmission dynamics, in turn, led to studies that aimed to control and prevent disease transmission.

Hib[5] was first identified as 'Pfeiffer's bacillus' in 1892 by Robert Pfeiffer, who succeeded in isolating and identifying the bacteria, but was unable to establish the diseases it caused. In the 1930s, Margaret Pittman (1901–1995), working at the Rockefeller Institute for Medical Research, was able to define the distinct strains (a–f) that differ in the composition of their polysaccharide capsules (Pittman, 1931, 1933).

At this point, Hib had gained its identity: Once the strains of the *Haemophilus influenzae* were defined, the bacteria (especially the b strain) were soon identified as the causative agent of *meningitis*, a potentially fatal inflammation of the brain tissue. Furthermore, the

4 The architects of germ theory, Louis Pasteur (1822–1895), Robert Koch (1843–1910) and Robert Pfeiffer (1858–1945), were able to formulate the logical conditions (known as Koch's postulates) needed to show that an organism 'x' is the cause of a disease 'y' (Bynum, 2006: 123).

5 The name *Haemophilus influenzae* refers to 'blood-loving' (hence Haemo [= blood] philus [= loving]) because of the requirement for blood factors for growth. *Influenzae* carries a false connotation of being the cause of the Spanish Flu in 1918 (e.g. in Wollstein, 1919).

bacteria were observed to circulate among young children, which increased the urge to provide protective measures against the disease. Understanding how the disease was transmitted between individuals was soon established. Upon successfully identifying the pathogen, and its mode of transmission, a mathematical theory on epidemics started to emerge. As Soper (1929: 34) emphasizes: 'Perhaps no events of human experience interest us so continuously, from generation to generation, as those which are, or seem to be periodic.' It is precisely this periodicity that led pioneers in the field of epidemiology to observe the patterns of transmission and formulate early models on them. For example, Kermack and McKendrick (1927) identified the basic categories of people (i.e. susceptible, infected and immune) observed in typical epidemics. Experimental epidemiology emerged as a platform for further investigation, expanding the focus of study towards understanding population dynamics of disease transmission (Amsterdamska, 2001). For mathematical epidemiologists, the quest was to locate 'global patterns of disease in time, space and population' (Fine, 1979). This search resulted in understanding the cyclic[6] nature of infection (Hamer, 1906), refining the mathematical theory of epidemics that endeavoured to identify which factors govern the 'spread of contagious epidemics' (Kermack and McKendrick, 1927). These public health concerns, however, required further discoveries on the microbiological level to develop preventive measures, such as vaccines.

Pittman's microbiological discoveries (1931, 1933) paved the way for the development of vaccines. However, it was only in the 1970s that polysaccharide vaccines began to be used in population-wide vaccination campaigns, and, to improve their efficacy, conjugate vaccines came to market in the late 1980s. Population-wide vaccination programmes using conjugate vaccines against Hib began in the US in 1985, in Finland in 1986, and in the UK in 1992. Hence, Hib vaccine development proceeded in two phases. The earlier polysaccharide vaccines were capable of protecting populations against the disease, but they did not diminish the transfer of the micro-organism, whereas conjugate vaccines did both. This, of course, as we will observe when modelling Hib transmission and circulation in a population, affects the estimates of the basic reproductive rate[7] and optimal herd-immunity threshold[8], which are important estimates for infectious disease epidemiology.

The birth of the fact—the identification of Hib and clarification of its severe disease forms—shows how knowledge produced in different fields, such as microbiology and mathematical epidemiology[9], are used in other domains, such as broadening the perspective of transmission among individuals to a population level. In a way, the singular fact (F1) gradually grew into the 'big fact' of population-level transmission.

6 Hamer's (1906) pioneering observations were on the cyclical nature of measles epidemics between the British and French populations.

7 The basic reproductive rate is defined as the average number of individuals infected by an infectious case during his or her entire infectious period, when he or she enters a totally susceptible population (Giesecke, 2002: 121).

8 Herd immunity threshold is the level of immunity in a population, which prevents epidemics (Giesecke, 2002: 124).

9 Mathematical epidemiology relied on developments in chemistry, especially by adopting the mass-action principle that considers mixing patterns of individuals as if they were 'gas molecules' (Hamer 1906).

To prevent and to protect: public health concerns on transmission dynamics

Transmission puzzled historians of medicine, especially in the early days when the distinction between infection and heredity was not well understood (Gaudillère and Löwy, 2001). In a simple classification of objects and patterns of transmission, *horizontal* is linked with infectious agents and *vertical* is seen as transmission of hereditary traits. But the story of infectious transmission is not that simple. Different infections are transmitted differently. Some are constantly present in a population, some occur in serious outbreaks but wane over time. The means of transmission may vary: some pathogens require physical contact and some are disseminated through the air. And it is sometimes not possible to specify a carrier, a susceptible, or an infected person, as some infectious agents do not always cause symptoms in their hosts. It seems that transmission is a stubborn piece of knowledge that requires thorough investigation in order to be incorporated into population-level transmission models.

In the case of Hib, it is transmitted from person to person via human excretions from the respiratory tract: by coughing, sneezing or kissing. However, Hib does not necessarily cause disease; the infected person may remain without symptoms,[10] capable of transmitting the bacteria but remaining healthy. This needs to be represented in dynamic Hib transmission models. Hib is therefore a good case study for explaining how the epidemiological features of a particular micro-organism may actually present challenges to mathematical representations of its spread. As described previously, the microbiological features of bacteria affect population-level transmission patterns. For example, an asymptomatic carrier is capable of spreading the pathogen, yet remains healthy. This is a challenge for effectively utilizing such knowledge.

In an effort to explain how disease transmission is related to population dynamics, I will discuss in detail the indirect population-level effects known as *herd immunity*. Herd immunity, according to Paul Fine (1993: 268), was based on the mass-action principle[11] that was initially formulated on the basis of a chemical property that says: 'the rate or velocity of a chemical reaction is a function of the product of the initial concentrations of reagents'. When applied to epidemiology, this means that changes in the number of susceptibles (infected and immunes) are expressed in successive steps. According to Fine (1993), herd immunity, as a concept, was developed in 1923 by Topley and Wilson, who studied infectious epidemics in laboratory mice and made herd immunity a topic of rigorous study (Fine, 1993: 266).

The theory of herd immunity claims that when diseases are passed from person to person, it is more difficult to maintain the chain of infection when large numbers of the population are immune. This can be calculated as a *herd immunity threshold*, which is the point at which the vaccinated percentage of a population is such that it effectively stops the spread of the infection, because there are no longer sufficient numbers of susceptibles to contract the disease. To calculate this threshold, one needs to define the basic reproductive number, R_0, which is the average number of directly infected individuals. In infections that are transmitted from person to person, the potential for spread is called the reproductive rate, which

10 An asymptomatic carrier.

11 Pioneering work by Hamer (1906) elaborated the mass-action principle in a deterministic model of measles outbreaks.

depends on the risk of transmission per contact and on the frequency of contacts. The reproductive rate is determined by:

- the probability of transmission in a contact between an infected individual and a susceptible individual
- the frequency of contacts in the population
- how long an infected person is infectious
- the proportion of the population that is already immune.

As specified, all these characteristics can be expressed in mathematical equations to provide numerical estimates of the transmission dynamics in a population (Giesecke, 2002). The next section explains in detail the challenge of mathematically modelling the general facts of disease transmission.

To summarize, through the birth and youth of disease transmission facts, the following details of Hib were defined: Hib bacteria is capable of causing life-threatening, invasive disease (such as meningitis, septicaemia or epiglottitis)—especially in young children. The severity of these conditions is a key factor motivating the quest for a detailed understanding of the dynamics of the infection in order to develop a means to control the spread of transmission. As briefly noted, infections caused by Hib have certain features that are challenging for the study of transmission. The natural locus of the bacteria is the human nasopharynx. This locus itself has two implications for transmission. First, it is difficult to define who is a carrier of the bacteria: there are no serological (blood) tests available, and taking a sample from the human nasopharynx is a much more complicated procedure than taking a simple blood sample. Second, transmission is easily spread through coughing, sneezing or close contact between a carrier and a susceptible. Moreover, an individual does not develop a life-long immunity to Hib. In fact, after infection there is only a short protective period before the individual is once more susceptible to re-infection. Hence, these general facts about Hib and its transmission in a population were established over many years of study on Hib diseases and vaccine development.

Thus, the *birth and youth* of facts shaped their identity. It seems that the facts of disease transmission largely led their lives in the mathematical expressions developed by early epidemiologists. The gradually built understanding of disease transmission encountered tension between the individual and the population, even in its infancy. Due to the complexity of transmission dynamics, the population-level effects of carriers or infected individuals can only be observed indirectly. These indirect observations are made by using transmission models that represent population-level dynamics. How these models mitigate the tension between the individual and the population is analysed in the following section.

Adulthood and (re)productive years: simulated 'facts' of transmission

This section explores how the set of general epidemiological facts (F1–F3), which form the *big fact* of population-level disease transmission, are expressed in probabilistic simulation models. In doing so, the aim of this section is to examine how facts about Hib transmission in a population are refined in models and to identify potential challenges encountered in this process. This section highlights the phase of *adulthood* in our life-history framework.

Adulthood is often related to maturity, stability and the potential for reproduction. The potential for reproduction is elaborated at greater length when these facts are further investigated in a simulation model.

Iterating facts of transmission in probabilistic models

Facts about Hib transmission were established in four models before being translated into the simulation model, which integrated the different aspects of the individual-based disease transmission process in Hib. Let us consider these translations as a process of stabilizing the accumulated knowledge claims of transmission.

The set of facts (F1–F3) are represented in the following four models:

M1: The *goodnight-kiss model* (Auranen *et al.*, 1996) provides insights into simple transmission dynamics within a closed population (e.g. a family). This model represents the fact that identifies Hib and its particular transmission pattern (F1).

M2: The *hierarchical Bayesian model to predict duration of immunity to Hib* (Auranen *et al.*, 1999) estimates declines in antibody concentration in order to predict the duration of immunity to subclinical Hib infection and serious invasive Hib disease. This model clarifies how immunity to Hib is acquired and maintained in a population (F2).

M3: The *dynamics of natural immunity* (Leino *et al.*, 2000) estimates the force of infection given different parameters. This model permits herd immunity effect estimates (F3).

M4: The *model on immunizing infections of Hib and cross-reactive antigens* (Leino *et al.*, 2002) is used to explain the differences in pre-vaccination incidence and age-distribution of invasive disease in different countries. This model refines factors that affect immunity (F2).

The main aim of the modelling exercise was not to develop further understanding of disease transmission as such, since it was well documented in the epidemiological literature. Instead, the aim was to assess the preventive and interventionist measures in the models. In other words, the aim was to understand the *prevention* of transmission and how it is reached by defining adequate immunity levels that result naturally (after the disease) or from vaccinations. This resulted in a 'need to *plan and evaluate different vaccination strategies*' (Auranen, 1999: 9). This need seems to be the reason why transmission is studied through investigations of immunity. Yet the main aim was accomplished via smaller steps (see Mattila, 2006a).

In developing the good-night kiss model, the aim was to study 'the transmission in appropriate subpopulations as well as to assess the relative importance of subpopulation and population transmission' (Auranen *et al.*, 1996: 2235). In other words, the study examined transmission in terms of estimating transmission rates within a family and a community in relation to Hib infection.

Focusing on how population-level disease transmission evolves required that the model be built to estimate family and community transmission rates of Hib (M1), which reveals an important characteristic of *maturation of facts*. The relatively general facts about transmission (F1–F3) become more detailed when new knowledge about the carriage of Hib is added. The model affected the facts; it incorporated predictions on the prevalence and

incidence of Hib carriage as a function of family size and age structure (Auranen, 1996: 2251). This means that what initially began as general features of transmission were later *narrowed down* and *sharpened* to show us that families with children of certain age-groups (those in day-care or at school) are more likely to harbour the infection, which also explains the apparently arbitrary occurrence of the infection in adult populations. The transmission happens within the family through casual physical contact, such as goodnight kisses (hence the model that bears this name).

The model to predict the *duration* of immunity to Hib (M2) established the decline in antibody concentration. It revealed the fact that the dynamics related to antibody concentration have important implications in predicting the consequences of different vaccination programmes (Auranen, 1999). The prevention of transmission by polysaccharide vaccines was observed by establishing the decline rate of antibodies in the model. Interestingly, the applicability of these observations extends to other bacterial infections with *similar antibodies* (i.e. pneumococcal and meningococcal infections). Hence, the facts of transmission became generalizable to other infections.

The dynamics of natural immunity model (M3) became an icon for later estimations of different levels of force of infection and evaluations of their impact on the duration of natural immunity. The principal idea was to explain the increase in invasive disease in unvaccinated cohorts. However, the estimations of the force of infection were useful because they helped parameterization of an agent-based model that simulated transmission. Once again, the exploration of natural immunity aimed at 'optimizing vaccination strategies', which was claimed to be the major aim of infectious disease epidemiology (Leino *et al.*, 2000: 583).

Furthermore, this general aim was translated into two specific steps (in M4): minimizing incidence of invasive Hib disease at a population level, and diminishing the number of colonizations (i.e. those environments where Hib protective antibody levels are below the estimate) by vaccination. This led to a model-based observation of transmission: there is a connection between the rate of decline of antibodies and the force of infection. This connection may have implications for population-level immunity (so called herd immunity). On the one hand, if the force of infection is low, vaccination may decrease the circulation of Hib bacteria and hence strengthen the herd immunity. On the other hand, if the force of infection is high, the vaccination may cause incidence of invasive disease in unvaccinated cohorts. Hence, transmission itself is altered, which has implications for vaccination planning (Leino *et al.*, 2000: 589). The model-based observation of transmission shows us how facts were iterated[12] from general epidemiological facts, which were established within theoretical and experimental epidemiology to model-specific facts, which highlight the specificities of transmission dynamics, carriage and immunity within a population.

Multiple interpretations of transmission

The practice of expressing the cluster of Hib transmission facts in an individual-based simulation model on transmission, immunity, and disease is a very detailed process in which the

12 Modelling is described as an iterative process within which the realistic description of the phenomenon under study is balanced with the amount of information available for the study (Auranen, 1999: 16).

general facts are simplified through a process of integration. This process, in which facts about transmission constitute a special mode of mathematization in terms of dynamic modelling, is typically in a mode of question-oriented modelling practices (Mattila, 2006a, 2006b). This model is capable of capturing agent-based dynamics, addressing mixing patterns and including vaccination effects on the circulation and transmission of Hib. As a fine-grained, population-simulation model, it is thus capable of describing the full notion of Hib transmission, immunity and disease (Auranen *et al.*, 2004). The ethnographic data on building, calibrating and testing the simulation model broadens the perspective on how facts about disease transmission are implemented within the simulation model.

The basic aim of the modelling activity was to come up with a transmission model. This task is neither obvious nor easy, even though it ultimately constitutes only *one part* of the more complex model. For those working on this problem, there is no guarantee that the process will succeed. As shown, population-level disease transmission is a manifold concept. For the purposes of modelling activities, it was disentangled, unpacked and, turned into smaller pieces with the help of auxiliary tools (e.g. representations of transmission patterns, estimates of natural immunity, etc.). One of these tools is called *force of infection* (defined as the rate at which susceptible individuals become infected, notated with 'λ'). Auxiliary tools are essential, as they force the process of mathematization to take into account the contextual specificities, such as type of infection, population structure and age dependency. Beneath the transmission model there is a population model. But turning this into a dynamic, growing population implemented with an S–I–S mechanism (Susceptible–Infected–Susceptible) turned out to be problematic. When the modellers encountered a bug in the population simulation programme, they chose to 'freeze' the population (that is, to model it *without* growth dynamics). The S–I–S pattern was nonetheless introduced into the frozen population, but the S–I–S pattern was, at this phase, limited only to carriage (immunity was forced artificially to result from the disease). A month later, in order to specify the infection caused by Hib and infection caused by other bacteria, exposure to cross-reactive bacteria was added to the model. Exposure to cross-reactives is likely to boost an individual's immunity. This interferes with the short period of immunity, caused by Hib (which will gradually turn back to susceptibility), as the S–I–S mechanism shows.

The interactional data show that, despite the long-term efforts to house facts about transmission within the models, it is still not quite clear to everyone what transmission actually *is*. The following extract, derived from my ethnographic observations, underlines this problem:

> In late 2002,[13] a computer scientist posed a crucial question in the meeting: 'What do we actually mean by transmission?' The answer reveals the usage of auxiliary concepts in order to reach the fact itself.
>
> *Statistician*: Well it means that we calculate through all contact sites according to the simulation model. And then we look, one individual at time, as we do in that [the model]. What is the force of infection with this individual? We have thought that it is the same all the time, the average force of infection. This is something we can calculate

13 Meeting at the National Public Health Institute, 17 November 2002 (001/22:35). Statistician is a post doc researcher, computer scientist is a senior researcher and epidemiologist (junior) is a PhD student.

from the data [we have]. And I have tried that and used the best parameter values I have had chance to estimate, and it seems to work quite well.

Computer scientist: Okay.

Statistician: So, in my opinion, I was sort of convinced that this works well, this transmission model. That we actually have the transmission model. And we have good parameter values and then we feed in/put in/implement CR [cross-reactives] as 'background radiation', on a stable, constant level, and it produces this disease.

Epidemiologist (junior): . . . and then there would be the variation required. That you have a constant CR and added Hib, as it is.

Statistician: Yes. But then among adults it means that it is only CR all the immunity we have, because Hib is reduced to so small a level.

Thus, it seems that facts about transmission need to be actively accommodated in the model. Be that as it may, as *facts* they also reveal a surprising degree of sensitivity: just a couple of months later, in early 2003, the testing and calibration activities were resumed. The transmission model was now subjected to preliminary evaluation by comparing the model-based results with the existing (real-world) knowledge on Hib transmission as documented in the literature. And these comparisons again revealed new sides to the facts: they were sensitive to the population in which they occurred. So even though the discussion was limited to Hib transmission, factors such as different vaccination schedules, different age structures in day-care and school groups, and different levels of CR, varied depending on the country or region in which the study was conducted.

Therefore, it seems that the facts of transmission do not simply reproduce themselves. To do so, it would be necessary to rigorously mimic the trends of the Finnish population, the vaccination plan and the CR exposure on which the model was based. This was discussed in terms of how well the model explains the known specificities of transmission in different countries (in Gambia and in the UK).

Epidemiologist (junior): [14] . . . so when we talk about these day-care groups and so on in relation to transmission, it is all relevant. So what they try to say is that in the developing countries, like in Gambia, from where we have the 'carriage data', all transmission has happened before the day-care age. But yet these data are not enough. So the study showed that the carriage is higher than here already during the first and second year of age.. . . And we also thought in the comparison between Finland and England that the disease occurs earlier than with us and . . . we sort of explained that with CR.

As expressed in this discourse, determining transmission patterns in a specific population requires modelling practices that elaborate transmission dynamics in detail. However, if data are scarce or absent, or if behavioural and mixing patterns are not clearly documented, mathematization may be difficult, and transmission dynamics can only be explored at a more general level. This, in the end, may diminish the usefulness of the modelled evidence (or leave unintended uncertainty in the models). Some levels of maturity were reached in the models, even though they were tailored to represent only the Finnish population. The

14 Meeting, 7 March 2003 (001/58:07).

challenge of acknowledging a shared interpretation of the transmission facts reminds us that tensions within scientific work may not always originate from social dynamics or division of labour (Mattila, 2006c). Often discrepancies have their roots in epistemological ground, as this section showed.

Senior facts: the dual nature of disease transmission

The analogy with old age captures two aspects of the life of facts: either the facts reach a respected position or they are retired. In common terms, we might think of the elderly reverentially, as repositories of accumulated wisdom. In this section, the senior facts will be discussed by reflecting on the dual nature of disease transmission.

The dual nature of transmission addresses the phenomenon as a threat and a promise of prevention. It is a threat that must be well understood in order to carry out vaccine development, strategic planning and protection. Yet it is also a promise that public health can protect using preventive interventions. This duality may remain invisible to the public. Yet, for public health authorities, facts of transmission are *senior facts*, respected for their information content, and their capability of being a reference point in decision-making. Let us return to the case of Hib and the transmission facts (F1–F3) established by the four models. As discussed previously, the models are primarily concerned with the prevention of transmission and, therefore, the general facts of transmission of and immunity to Hib. These models emphasize the dual nature of disease transmission. Even though public health measures aim primarily to prevent transmission, transmission is also the force that induces immunity in the population. In order to optimize preventive measures (vaccinations), one needs to take into consideration possible decreases in natural immunity that result from the circulation of the bacteria.

Thus, these models elaborate, represent and reproduce facts about immunity—an immunity that protects against and prevents transmission. Each of these models tells a story of its own, but together they contribute to the development of realistic estimates to prevent transmission. In a way, they reveal the dual nature of transmission that must be fully understood. Transmission does not only pose a threat of severe disease, it also has a protective purpose, and interfering with this cycle (by vaccinations) may alter the delicate balance and cause new cases of disease in older age groups. Hence, this aspect of transmission clearly underlines the challenges of drawing firm conclusions about disease transmission and representing them in models. It also reminds us of the public health perspective, discussed earlier: it is not enough to explain the singular causes of disease, as it is necessary to expand the explanation to cover the causes and effects of disease transmission in a population.

Another way of looking at the obsolescence of facts is by reflecting on the unequal distribution of Hib vaccinations in national vaccination programmes. Thus, even after Western societies reduced the problem of disease transmission, it still continues to be a threat in many parts of the world. Simulated facts encourage us to keep the problem in mind, and urge us to develop tools for public health advisory work on reshaping vaccination policies. For example, the implementation of Hib vaccination in England and Wales can be considered a public health success story. The declining trend in Hib cases is indicative of the efficacy of the vaccine. Why call it a success story? To recap, Hib is the leading cause of

bacterial meningitis, and other invasive diseases, such as pneumonia and epiglottitis. Worldwide, there are still approximately 3 million cases of serious disease each year, with approximately 386,000 Hib-related deaths in infants aged 4–18, and bacterial meningitis represents 52 percent of these infections. Societies continuously struggling with Hib diseases are in the developing world (WHO, 2006 445–452). Simulation models can be used as administrative tools to facilitate the alleviation of the disease burden in these countries.

The main use of simulation models is to help optimize vaccination strategies in the developing world. This was one of the aims of the Helsinki modelling project: to create a general tool for a variety of purposes in order to overcome the problem of unavailable surveillance or serological data. But the challenge remains to persuade developing countries to include Hib vaccinations in their strategies. In pursuit of these aims, simulation models may also function as flexible tools to carry the *senior* facts of transmission to other domains. In other words, as expressed in Leino (2003: 10):

> In the developing world data on the disease burden are limited and although WHO and GAVI[15] advocate Hib conjugate vaccination, the major question remains whether universal vaccination will be at all feasible in the poorest economies. [...] Ultimately information on Hib transmission allows predictions to be made of the effects on the population level following large-scale vaccinations.

This is clearly a strong reminder that a fact (about Hib transmission) is capable of adopting an expert position. It also explains the rationale behind using models to refine the facts, and to overcome possible shortcomings in available data. The modelled facts regarding decreases in transmission, the vaccination coverage needed for an optimal herd immunity level, and the falling number of severe cases should convince decision-makers of the necessity of vaccinations. It is through the *seniority* of these facts and their generalizability, despite the fine details of their modelled context, that they are imbued with sufficient persuasive power to fight against superannuation. Nonetheless, striking a balance remains an open-ended challenge. In summary: the characterization of the *seniority* of facts allows old facts to hold respected positions. With regard to the question of disease transmission, public health authorities seek advice and evidence for decision-making processes from senior facts, from knowledge claims that contain information applicable to different contexts.

Discussion: understanding disease transmission through the *life history* of facts

As a conceptual framework, what does *life history* highlight? Briefly, in each phase, knowledge claims—facts—are refined, sharpened, circulated and represented in a particular way, allowing them to carry on with their lives. Mathematical representations and models function as a key to these heterogeneous practices. From early observations of transmission in populations to fine-grained simulation models, the facts of transmission dynamics are expressed in mathematical terms. Life history, as an analytical framework, not only shapes the linear timeline of the developmental process, it also provides insights into how facts live

15 The Global Alliance for Vaccinations and Immunization.

their lives. *Birth and youth* emphasizes the gradual nature of the process, within which generally accepted epidemiological facts emerge and become accepted within research communities. Maturation towards *seniority* is expressed in terms of generalizability, which refers to the potential to broaden the epistemic scope of available knowledge. *Life history*, furthermore, invites us to explore the *vita activa* of facts in a more metaphorical sense: it shows the flux embedded in the continuum from knowledge production to dissemination. It seems that the idea of knowledge production, which is well documented and discussed, has led us to believe that once knowledge or facts are produced they are simply 'out there', ready to be used (e.g. Knorr Cetina, 1981; Latour and Woolgar, 1986 [1979]). Yet, *vita activa* moves beyond these assumptions and reveals the continuous effort required to understand multifaceted disease transmission phenomena by expressing them in mathematical representations.

Life history, as a conceptual framework, is developed through the utilization of multiple sources of data. Thus, it facilitates the analysis of micro-level interactions between modelling practices and the more general public health studies of disease transmission. Hence, the parallel development of early theories on transmission dynamics, the difficulties encountered in first identifying the specific bacterial agent that causes *Haemophilus influenzae*, and growing awareness of population-level dynamics tell a different story. Through the life history of facts, some moments in the developmental process of epidemiological knowledge were traced and the interdependency between epidemiological knowledge and mathematical representations was established.

Acknowledgements

This research was conducted as part of a Leverhulme Trust/ESRC funded research project 'Nature of Evidence: How Well Do ''Facts'' Travel? (Grant number F/07004/2) at the Department of Economic History, London School of Economics and Political Science. I wish to thank Professor Mary Morgan and Dr Patrick Wallis, and my colleagues in the 'Facts' project, for their feedback on the previous versions of the manuscript. I thank the anonymous referees for their comments. Earlier versions of this article were presented at the British Academy Conference 'Enquiry, Evidence and Facts' (December 2007). I wish to also thank my colleagues at the conference for their insightful discussion. I am grateful for the collaboration with researchers at the National Public Health Institute, Helsinki and the University of Helsinki, Department of Mathematics and Statistics.

References

Amsterdamska, O. (2001). Standardizing epidemics: Infection, inheritance and environment in pre-war experimental epidemiology. In J.-P. Gaudillère, & I. Löwy (Eds), *Heredity and infection: The history of disease transmission*. London: Routledge.

Auranen, K. (1999). *On Bayesian modelling of recurrent infections*. Rolf Nevanlinna Research Report 26, University of Helsinki.

Auranen, K., Ranta, J., Takala, A., & Arjas, E. (1996). A statistical model of transmission of Hib bacteria in a family. *Statistics in Medicine, 15*, 2235–2252.

Auranen, K., Eichner, M., Käyhty, H., Takala, A., & Arjas, E. (1999). A hierarchical Bayesian model to predict the duration of immunity to Hib. *Biometrics, 55*(4), 1306–1314.

Auranen, K., Eichner, M., Leino, T., Takala, A., Mäkelä, P. H., & Takala, T. (2004). Modelling transmission, immunity and disease of Haemophilus influenzae type b in a structured population. *Epidemiology and Infection, 132*, 947–957.

Becker, H.S. (2007). *Telling about society*. Chicago: U Chicago Press.

Bynum, W.F. (2006). *The Western medical tradition: 1800–2000*. Cambridge: Cambridge UP.

Creager, A. (2002). *The life of a virus*. Chicago: U Chicago Press.

Daston, L. (2000). Introduction: The coming into being of scientific objects. In L. Daston (Ed.), *Biographies of scientific objects*. Chicago: U Chicago Press.

Fine, P. (1979). John Brownlee and the measurement of infectiousness: A historical study in epidemic theory. *Journal of the Royal Statistical Society. Series A (General), 142*, 347–362.

Fine, P. (1993). Herd immunity: History, theory, practice. *Epidemiologic Reviews, 15*, 265–302.

Gaudillère, J.-P., & Löwy, I. (2001). *Heredity and infection: The history of disease transmission*. London: Routledge.

Giesecke, J. (2002). *Modern infectious disease epidemiology*, 2nd edn. London: Arnold.

Hamer, W.H. (1906). Epidemic disease in England. *The Lancet i*, 733–739.

Kermack, W.O., & McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 115* (772), 700–721.

Knorr Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Oxford: Pergamon Press.

Latour, B., & Woolgar, S. (1986 [1979]). *Laboratory life: The social construction of scientific facts*. London: SAGE.

Leino, T. (2003). *Population immunity to Haemophilus influenzae type b—Before and after conjugate vaccines*. Publications of the National Public Health Institute A23/2003. Helsinki: Hakapaino.

Leino, T., Auranen, K., Mäkelä, P.H., & Takala, A. (2000). Dynamics of natural immunity caused by subclinical infections: Case study on Haemophilus influenzae type b (Hib). *Epidemiology and Infection, 125*, 583–591.

Leino, T., Auranen, K., Mäkelä, P.H., Käyhty, H., Ramsey, M., Slack, M. *et al.* (2002). Haemophilus influenzae type b and cross-reactive antigens in natural Hib infection dynamics: Modelling in two populations. *Epidemiology and Infection, 129*, 73–83.

Leonelli, S. (forthcoming). Packaging data for re-use: Databases in model organism biology. In M. Morgan, & W.P. Howlett (Eds), *How well do 'facts' travel?* Cambridge: Cambridge UP.

Mansnerus, E. (forthcoming). Using models to keep us healthy: Productive journeys of facts across public health networks. In M. Morgan, & W.P. Howlett (Eds), *How well do 'facts' travel?* Cambridge: Cambridge UP.

Mattila, E. (2006a). Interdisciplinarity in the making: Modelling infectious diseases. *Perspectives on Science: Historical, Philosophical, Sociological, 13*, 531–553.

Mattila, E. (2006b). *Questions to artificial nature: A philosophical study of interdisciplinary models and their functions in scientific practice*. Philosophical Studies from the University of Helsinki, no. 14. Helsinki: Hakapaino.

Mattila, E. (2006c). Struggle between specificity and generality: How do infectious disease models become a simulation platform? *Simulation: Pragmatic Constructions of Reality—Sociology of the Sciences Yearbook 25*, 125–138.

Morgan, M. (forthcoming). 'Introduction'. In M. Morgan, & W.P. Howlett (Eds), *How well do 'facts' travel?* Cambridge: Cambridge UP.

Mendelsohn, A.J. (2003). Lives of the cell. *Journal of the History of Biology, 36*, 1–37.

Pittman, M. (1931). Variation and type specificity in the bacterial species Haemophilus influenzae. *Journal of Experimental Medicine, 53*, 471–492.

Pittman, M. (1933). The action of type-specific Haemophilus influenzae antiserum. *Journal of Experimental Medicine, 58*, 683–706.

Rheinberger, H.-J. (2000). Cytoplasmic particles. In L. Daston (Ed.), *Biographies of scientific objects*. Chicago: U Chicago Press.

Soper, H.E. (1929). The interpretation of periodicity in disease prevalence. *Journal of Royal Statistical Society, 92*, 34–72.

WHO (2006). *Weekly epidemiological record*. 24 November, 47(81), 445–452.

Wollstein, M. (1919). Pfeiffer's bacillus and influenza: A serological study. *Journal of Experimental Medicine, 30*, 555–568.