REVIEW 10.1111/1469-0691.12183

Contact investigations for outbreaks of *Mycobacterium tuberculosis*: advances through whole genome sequencing

T. M. Walker[1], P. Monk[2], E. Grace Smith[3] and T. E. A. Peto[1,4]

1) *Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford,* 2) *East Midlands South Health Protection Unit, Leicester,* 3) *West Midlands Public Health Laboratory, Health Protection Agency, Heart of England NHS Foundation Trust, Birmingham and* 4) *Oxford National Institute of Health Research Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK*

Abstract

The control of tuberculosis depends on the identification and treatment of infectious patients and their contacts, who are currently identified through a combined approach of genotyping and epidemiological investigation. However, epidemiological data are often challenging to obtain, and genotyping data are difficult to interpret without them. Whole genome sequencing (WGS) technology is increasingly affordable, and offers the prospect of identifying plausible transmission events between patients without prior recourse to epidemiological data. We discuss the current approaches to tuberculosis control, and how WGS might advance public health efforts in the future.

Keywords: Contact investigation, outbreak, tuberculosis, whole genome sequencing

Clin Microbiol Infect

Corresponding author: T. M. Walker, Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Headley Way, Headington, Oxford, OX3 9DU, UK
E-mail: timothy.walker@ndm.ox.ac.uk

Introduction

The decline in tuberculosis incidence and mortality in western Europe since the mid-18th century pre-dates the discovery of the tubercle bacillus in 1882 and the development of drug treatments in the 1940s. The reasons for this decline are disputed, but hypotheses range from improvements in living standards to the isolation of 'consumptives' in Poor Law infirmaries and sanatoria. By 1990, this trend had been reversed [1].

Historical trends in Africa, Asia and South America are less well characterized, but historical and phylogeographical data are consistent with the epidemics on these continents dating back to the late 19th century, after the disease was probably (re)-introduced by European colonizers [2,3]. Although this is relatively late into the colonial period, in India the timing coincides with a surge in British troop numbers after the 1857 mutiny, and the building of the railways that provided efficient channels of transmission for disease [4]. The global burden of disease is now felt most acutely on these continents, where many of the world's 2 billion people infected with latent or active tuberculosis can be found [5].

Today, tuberculosis remains a disease of poverty in high-income and low/middle-income countries alike. Without major breakthroughs among experimental vaccines [6], available control measures include contact tracing, active case-finding, prophylaxis, and treatment. In high-income countries, contact investigations have benefited from advances in genotyping techniques over the past two decades. The arrival of rapid-turn-around whole genome sequencing (WGS) technology has the potential to guide public health teams in all settings with unprecedented precision.

Epidemiology

Observations that patients with pulmonary tuberculosis often do not lead to any secondary cases fuelled debate in the 19th century about whether the disease was communicable at all [7]. Although this issue was definitively settled by Koch's discovery, how a disease with a predominance of

non-infectious hosts has managed to infect one in three individuals on the planet remains poorly understood. Patients with latent tuberculosis infection have an expected 10% lifetime risk of reactivation (this rises to 10% per year if the patients are infected with human immunodeficiency virus) [5]. Among patients with active tuberculosis, approximately half have pulmonary disease; half of these are sputum smear-positive [8] and are hence considered to be infectious. In a meta-analysis of pooled data from 41 studies, the risk of infection among household contacts of these patients with 'open' tuberculosis has been quantified at 50% for the development of latent tuberculosis infection and <5% for the development of active disease [9]. Thus, on average, each patient with open tuberculosis must have the unlikely equivalent of >20 household contacts to result in one further infectious case (Fig. 1). Reports that hyperinfectious individuals can be responsible for a large amount of secondary cases in community outbreaks [10–13] and in experimental settings [14,15] may offer a potential explanation. Indeed, mathematical modelling has predicted that if the success of tuberculosis can be attributed to 'super-spreaders', their identification and treatment will be key not only to the control of outbreaks, but also to combating the disease as a whole [16]. However, the degree to which super-spreaders account for transmission in any given community has so far been difficult to quantify.

Public Health Control Measures

Mobile mass X-ray screening was introduced as a tuberculosis control measure in industrialized countries in the 1930s. By the 1970s, a realization that most patients with active tuberculosis seek healthcare for their symptoms led to the phasing out of mass screening and a greater focus on diagnostic services [17]. Although screening remains relevant among patients who are less likely to seek healthcare [18], targeted contact investigations to identify 'source' and 'secondary' cases within outbreaks are now standard public health practice. Guidance varies across Europe and the USA, with some countries initiating contact investigations only for potential 'source cases' (patients with smear-positive pulmonary tuberculosis) and others recommending contact investigations for 'index cases' in general, regardless of whether they are considered to be plausibly infectious [19,20]. The standard model for contact investigations has been to trace potentially exposed individuals across widening 'concentric circles' until the rate of positive screening test results reflects the background community prevalence of disease [21]. Most contact investigations focus on household contacts first, and are extended into the wider community only if at-risk individuals are identified or if a wider outbreak is suspected. These environments include schools and workplaces, both of which are relatively structured settings in which to conduct contact investigations, but also pubs/bars or homeless shelters, where attendees are more transient [18]. Investigations are dependent on the contacts being named by an index case and the proportion of 'close contacts' that screen positive for latent or active disease on initial investigation. Because patients from some of the social groups at highest risk of tuberculosis may not know the names of their contacts or may be reluctant to volunteer names, owing to social stigma or concerns about the legal implications of naming associates, this approach has its limitations [18,22,23].

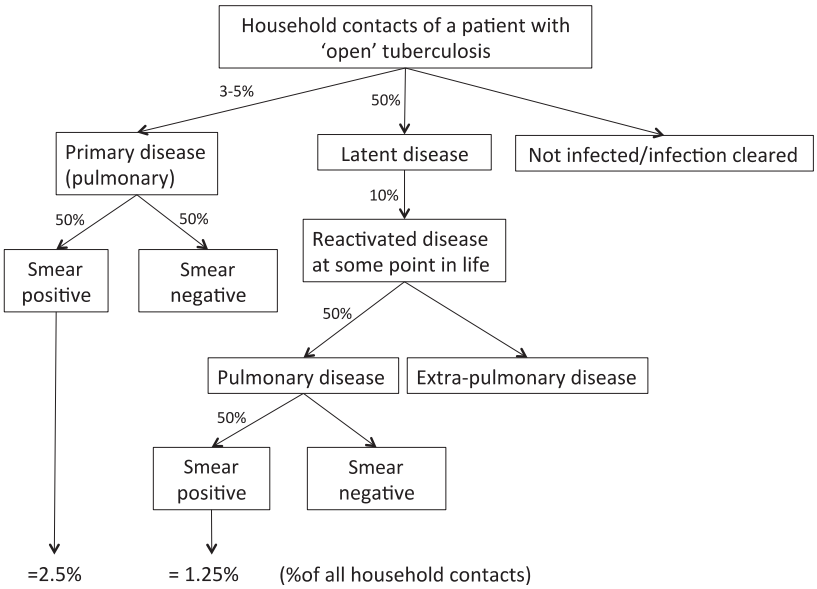

FIG. 1. Proportion of household contacts likely to develop the infectious form of the disease.