Pulmonary Perspective

Immune Reconstitution and “Unmasking” of Tuberculosis during Antiretroviral Therapy

Stephen D. Lawn1,2, Robert J. Wilkinson3,4,5, Marc C. I. Lipman6, and Robin Wood1

1The Desmond Tutu HIV Centre, Institute for Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa; 2Clinical Research Unit, Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom; 3Mycobacterial Immunology, Institute of Infectious Disease and Molecular Medicine, and Department of Medicine, University of Cape Town, Cape Town, South Africa; 4Division of Medicine, Imperial College London, London, United Kingdom; 5National Institute for Medical Research, London, United Kingdom; and 6Department of HIV Medicine, Royal Free Hospital, London, United Kingdom

Tuberculosis (TB) is the most common opportunistic disease in HIV-infected patients during the initial months of antiretroviral therapy (ART) and presents a great challenge to ART programs in resource-limited settings. The mechanisms underlying development of TB in this period are complex. Some cases may represent progression of undiagnosed subclinical disease present before starting ART, emphasizing the importance of careful screening strategies for TB. It has been suggested that progression in such cases is due to immune reconstitution disease—a phenomenon in which dysregulated restoration of pathogen-specific immune responses triggers the presentation of subclinical disease. However, whereas some cases have exaggerated or overtly inflammatory manifestations consistent with existing case definitions for IRD, many others do not. Moreover, since ART-induced immune recovery is a time-dependent process, active TB may develop as a consequence of persisting immunodeficiency. All these mechanisms are likely to be important, representing a spectrum of complex interactions between mycobacterial burden and changing host immune response. We propose that the potential range of effects of ART includes (1) shortening of the time for subclinical TB to become symptomatic (a phenomenon often referred to as “unmasking”), (2) increased rapidity of initial onset of TB symptoms, and (3) heightened intensity of clinical manifestations. We suggest that the term “ART-associated TB” be used to refer collectively to all cases of TB presenting during ART and that “immune reconstitution disease” be used to refer to the subset of ART-associated TB cases in which the effect on disease severity results in exaggerated and overtly inflammatory disease.

Keywords: HIV; tuberculosis; antiretroviral; immune reconstitution; IRIS

Since the mid-1990s, availability of antiretroviral therapy (ART) has transformed the prognosis of HIV-infected people living in high-income countries and more recently in resource-limited settings (1, 2). Through suppression of HIV replication, ART permits both quantitative and functional reconstitution of the immune system (3, 4), thereby reducing the risk of new opportunistic infections and neoplasia (1, 5). However, an appreciable (albeit diminished) risk of opportunistic infection persists during ART, especially during the first few months of treatment. Tuberculosis (TB) is the most frequent of these both in high-income and resource-limited settings (6–15).

Although heightened clinical vigilance may, in part, account for increased ascertainment of cases of TB during the initial months of ART, several different mechanisms may also underlie the temporal distribution of cases. ART-induced immune recovery occurs gradually over time and so active TB may develop in some patients as a consequence of persisting immunodeficiency. Conversely, restoration of pathogen-specific immune responses during treatment may cause subclinical TB to manifest or be “unmasked” (16–19). Of these latter cases, there is debate as to whether some or all of these should be termed “immune reconstitution disease” (IRD); clarity on this issue is needed.

The burden of TB during early ART is particularly great in resource-limited settings where this presents a major challenge in treatment programs (9). Management of these two diseases requires concurrent use of two multidrug regimens, with associated high pill burden, overlapping toxicity, and pharmacokinetic interactions (20). Furthermore, patients with incident TB during ART have increased mortality risk and may be a source of TB transmission within ART clinics (9, 21). Understanding the underlying mechanisms will facilitate development of preventive and therapeutic strategies. In this article, we consider how ART-induced immune recovery is likely to mediate a range of effects on both the timing and manifestations of active TB.

BURDEN OF TB DURING EARLY ANTIRETROVIRAL THERAPY

A number of reports have described the incidence of TB in patients receiving ART in high-income countries (12–15). The largest of these is a study of European and North American cohorts that included over 17,000 patients with a median baseline CD4 cell count of 280 cells/μl (12). The prevalence of TB at baseline was not known but the incidence rate was 1.3 cases per 100 person-years during the first 3 months of ART (12). Beyond this time point, the risk of TB decreased sharply (Figure 1A).

As might be expected, incidence rates of TB are substantially higher in ART cohorts in sub-Saharan Africa and other resource-limited settings (7, 9–11, 22). For example, in a community-based ART cohort in South Africa with a median baseline CD4 cell count of 96 cells/μl, prevalent TB (disease either newly diagnosed or in patients already receiving treatment) was reported among 25% of all patients at baseline (9). Despite the
high yield of prevalent TB detected during pre-ART screening of all patients with suggestive symptoms, the rate of incident TB in the first 3 months of ART was nevertheless extremely high at 23 cases per 100 person-years (Figure 1B). All cases had new onset of symptoms after initiation of ART and therefore represented true incident disease. The risk of TB decreased steeply during the first year of ART (Figure 1B) and was strongly associated with improving immune status (9).

**ROLE OF IMMUNODEFICIENCY**

Patients accessing ART programs in resource-limited settings typically have advanced immunodeficiency (2) and therefore have very high TB incidence rates just before commencing ART. An exceptionally high rate in excess of 50 cases per 100 person-years has been reported in African patients with World Health Organization clinical stages 3 and 4 disease in low-TB-burden countries in Europe and North America (data are from Reference 12) and (6) a community-based ART program in Cape Town, South Africa, where the burden of TB is very high (data are from Reference 9). Power trend lines are shown. PY = person-years.

**Figure 1.** Graphs showing the high incidence of tuberculosis (TB) during first 3 months of antiretroviral treatment (ART) and the subsequent rapid decrease. Data are from (6) cohorts in low-TB-burden countries in Europe and North America (data are from Reference 12) and (6) a community-based ART program in Cape Town, South Africa, where the burden of TB is very high (data are from Reference 9). Power trend lines are shown. PY = person-years.

ROLE OF IMMUNODEFICIENCY

Patients accessing ART programs in resource-limited settings typically have advanced immunodeficiency (2) and therefore have very high TB incidence rates just before commencing ART. An exceptionally high rate in excess of 50 cases per 100 person-years has been reported in African patients with World Health Organization clinical stages 3 and 4 disease in the pre-ART era, for example (23). In cohorts in which incidence rates are so high, active but “subclinical” TB is likely to be present in a proportion of patients at any given point in time. This could either progress to symptomatic disease or remain undiagnosed until death of the patient.

Post mortem studies of patients who died with HIV/AIDS support the supposition that rates of subclinical TB are high in Africans with advanced HIV infection; occult disseminated TB has been detected in up to 54% of cadavers (24, 25). Furthermore, in studies in which HIV-infected patients in southern Africa were actively screened for TB, high rates of culture-proven pulmonary disease have been detected despite the absence of suggestive symptoms (26–29). In the absence of a standard definition for subclinical TB, the proportions reported by such studies will therefore vary according to the criteria used. However, the fact that the World Health Organization’s DOTS (directly observed treatment, short course) strategy places greatest emphasis on investigation of symptomatic patients is problematic and most subclinical disease is likely to remain undetected pre-ART.

The limited sensitivity of TB diagnostic tests when applied to patients with advanced immunodeficiency further compounds poor TB case-finding in patients preparing to start ART. Sputum smear microscopy, for example, which is the mainstay of TB diagnosis in much of the developing world, typically detects less than 50% of active pulmonary disease in those with HIV infection (9, 27, 30). Thus, a proportion of the TB presenting during early ART may simply reflect failure of diagnosis of suspected cases before treatment initiation. This burden will vary between settings according to the intensity of screening and diagnostic facilities available.

We suggest that subclinical TB among patients entering ART programs provides much of the reservoir of infection that fuels the high rates of symptomatic TB presenting during early ART. Because ART-induced immune recovery is time dependent, some of these cases may progress to symptomatic disease during early ART in the context of persisting immunodeficiency. In addition, rates of latent *Mycobacterium tuberculosis* infection and new exogenous infection are also very high in high-TB-prevalence settings (31, 32); these may also progress and contribute to the burden of active disease during early ART.

**ROLE OF IRD**

Having discussed the role of immunodeficiency, we now consider IRD as a further important potential mechanism in the presentation of TB (33–35). IRD (also known as immune reconstitution inflammatory syndrome or IRIS) is believed to be caused by dysregulated recovery of pathogen-specific immune responses during the initial months of ART. This leads to development of unusual and overtly inflammatory disease, most commonly in association with chronic viral, mycobacterial, and invasive fungal infections (33–36).

TB-associated IRD presents in two principal ways. The most common form occurs in patients with TB receiving effective treatment who subsequently start ART; ART-induced immune recovery results in the deterioration of the clinical manifestations of TB. This has been reported to occur in 8 to 43% of patients with TB starting ART (35, 37–42). In the context of this review, however, the second form of TB IRD is the more relevant. Here, recovery of host immune responses triggers the clinical presentation or “unmasking” of previously subclinical TB during the early months of ART (33–35). It seems likely that these two types of IRD may be related phenomena, but immunologic studies to support this assertion are lacking.

Reports of TB IRD causally causing presentation of TB after ART initiation are relatively few (16, 18, 19, 43, 44), the clearest example being that reported by Goldsack and colleagues (19). In this report, a 41-year-old African with advanced HIV in London commenced ART. Two weeks later, he presented again with rapid-onset fever, breathlessness, and hypoxia. Although his chest radiograph was normal before treatment, a repeat radiograph revealed florid miliary shadowing that was microbiologically proven to be TB. The onset of this complication was temporally associated with a rapid immunologic and virologic response to ART. The patient developed adult respiratory distress syndrome, requiring mechanical ventilatory support, but he subsequently had a good response to anti-TB drugs and adjunctive corticosteroid therapy. An appropriate diagnosis of TB...
IRD was made, which is entirely consistent with similar cases we have seen in South Africa.

Breen and colleagues in London, United Kingdom, described a series of 19 cases of incident TB during ART and hypothesized that the presentation of 13 of these cases may have been attributable to IRD (16). The rationale for this was largely based on the temporal distribution of cases; 13 early-onset cases were clustered within the first 4 months of ART and many of these had a marked inflammatory response and systemic illness. Also, a paradoxical reaction after initiation of anti-TB therapy was observed in eight (62%) of the early cases compared with none of the late-onset group (16).

In agreement with the report by Breen and colleagues, workers in Uganda speculated that IRD was also the cause of the large number of new cases of TB diagnosed during ART in their setting (18). In a further report from the same country, all cases of incident TB during the first year of ART were described by the authors as being due to IRD (45). Park and coworkers reported from South Korea that 82% of new TB events occurring during the first year of ART (either new TB cases or paradoxical reactions at a new disease site) were attributable to IRD (17).

Evaluation of the role of IRD among cases of incident TB during ART is hindered by the lack of a diagnostic test or a clinical case definition specific to this condition. Importantly, however, the existing literature indicates that, for a diagnosis of IRD to be made, either the clinical manifestations and/or the clinical course of disease should be atypical and be consistent with an exaggerated, overtly inflammatory host response (34, 35, 46). Our own clinical experience gained during several studies of TB during ART in South Africa (7, 9, 42), however, is that most cases of TB diagnosed in the initial months of ART have typical clinical presentations. Only a subset of such cases present with unusual, overtly inflammatory manifestations that would be consistent with a clear diagnosis of IRD. Moreover, in high-TB-burden settings, it cannot be reasoned that the timing of onset of TB during the first months of ART is particularly unusual. Thus, in the context of the existing literature on IRD, it would appear appropriate to use the term IRD to refer to some but not all cases of incident TB during the initial months of ART.

HYPOTHESIS: A SPECTRUM OF EFFECTS OF IMMUNE RECOVERY ON TB PRESENTATION

Thus far, we have considered contrasting reasons why TB may present during early ART, with TB developing in the context of persisting immunodeficiency or TB presenting as TB IRD. However, we suggest that these represent opposing ends of a spectrum of important effects of ART-induced immune recovery on TB presentation.

Because immunopathologic host responses to M. tuberculosis are central to the clinical presentation of TB, we suggest that development of symptoms at a given anatomic site is a function of two key parameters: the mycobacterial antigen burden and the intensity and quality of the associated host inflammatory response. It is the interrelationship between these that is likely to determine the threshold for symptom onset and to affect the timing and severity of cases of TB presenting during early ART.

The supposition that high mycobacterial load is a risk factor for TB IRD is strongly suggested by two observations: first, the risk of IRD is greater the earlier ART is started during anti-TB therapy when the residual mycobacterial burden is high; second, the risk of TB IRD is increased in those with disseminated TB (33, 35, 42). Considerable improvements in immune function occur during the initial weeks of ART in most ART-naive patients living in resource-limited settings, even in those with the lowest baseline CD4 cell counts (47–49). Thus, the potential for cases of TB presenting in this period to be modified by changes in immune function is substantial. We suggest that immune recovery has three principal effects on the progression of subclinical TB to symptomatic disease, namely the timing of onset, the rapidity of initial symptom onset, and the overall intensity of clinical manifestations.

Timing of Onset

In an HIV-infected patient with a certain level of immune function, symptoms of TB are likely to develop at a given anatomic site once the mycobacterial burden that is sufficient to trigger host inflammatory responses has been reached (Figure 2A). If ART is initiated, however, TB-specific immune function increases rapidly (49, 50), and this would have the effect of rapidly lowering the threshold for symptom development (Figure 2B). This process may be augmented by immune-mediated liberation of mycobacterial antigen, further enhancing immune responses. In patients with subclinical TB, this would cause TB to manifest much sooner than would otherwise have occurred in the absence of ART (Figure 2B). In effect, the initiation of ART may serve as a therapeutic challenge that unmasks subclinical disease, triggering the presentation of TB. This accelerated progression

![Figure 2](image-url)
to symptomatic TB would result in the temporal clustering of cases in the initial 3 months of ART as observed in Figure 1 and in the report by Breen and colleagues (16).

Rapidity of Onset
In patients with advanced immunodeficiency, TB often has a fairly insidious onset. In contrast, however, our clinical impression is that, in cases of TB presenting during the initial weeks of ART, the initial onset of symptoms is often unusually rapid. This is consistent with the speed with which antimycobacterial immune responses recover during early ART (49, 50).

Intensity of Clinical Manifestations
When immune recovery causes unmasking of subclinical TB, the severity of clinical manifestations at a given anatomic site is likely to depend on the interrelationship between two main variables: the mycobacterial antigen load and the rate, intensity, and quality of immune recovery (Figure 3). In those in whom immune function remains very poor, TB is likely to remain subclinical despite an increasing mycobacterial load. Those patients with an intermediate rate of immune recovery would tend to develop clinical disease with a relatively “normal” presentation. In contrast, we propose that patients with dysregulated immune recovery are the ones most likely to develop unusual and severe manifestations that characterize IRD. This would be most pronounced in those with a high mycobacterial load (Figure 3). Within this model, therefore, only a proportion of patients in whom ART causes the unmasking of TB in the initial months of ART would fulfill existing case definitions for IRD. With a disease that is as heterogeneous in its clinical presentation as TB, it will clearly be difficult to develop a robust case definition for these cases.

CLINICAL AND PROGRAMMATIC IMPLICATIONS
The high early burden of TB during ART highlights the fundamental problem of how to efficiently screen for TB among patients preparing for ART. The less adequate the pre-ART screening processes, the greater the burden of TB is likely to be after starting treatment. Optimal screening strategies need to be defined; it is possible that all patients entering ART programs in resource-limited settings should ideally undergo culture-based screening for TB regardless of the presence or absence of symptoms. The major need to upgrade laboratory facilities and expand availability of culture-based diagnosis in resource-limited settings has been recognized and more sensitive tools for TB diagnosis at the point of care are desperately needed (30).

Although screening for TB among patients with advanced HIV-associated immunodeficiency is difficult, ART may paradoxically serve as a useful diagnostic tool, serving to unmask disease. Research is needed to determine the relative merits of delaying initiation of ART to permit full culture-based investigation for TB versus starting ART without delay if there is no initial evidence of TB. Because few of the cases of unmasking TB have increased clinical severity and yet delays in ART initiation may be associated with high mortality risk in this setting (51), we would not advocate a strategy of deferring ART while waiting for TB culture results in all patients.

Many in the field agree that management of those with moderate or severe TB-associated IRD should include corticosteroids (33–35). Pending the results of a randomized controlled trial in South Africa, however, the only evidence for this comes from case reports and expert opinion. In view of the seemingly normal clinical presentations of the large majority of cases of unmasking TB presenting in the initial months of ART, however, it seems likely that most patients who develop TB during early ART are unlikely to require adjunctive immunomodulatory therapy. On the basis of the same premise, there may be little merit in a strategy of using corticosteroids as prophylaxis against unmasking TB in high-TB-burden settings. Potential benefits, if any, may well be restricted to a small minority of patients and could be outweighed by adverse events (52).

Finally, it has been speculated that the unmasking of TB during the early months of ART may contribute to the overall burden of TB in resource-limited countries where ART is being rolled out (16). This is of course true and is an inevitable consequence of hundreds of thousands of lives being saved by expanding access to ART. However, the longer term burden of TB during ART (rather than that occurring in the initial months of treatment) is much more important with regard to TB control at the population level (9, 49, 53).

CONCLUSIONS AND SUGGESTED TERMINOLOGY
Using data and clinical experience from both low- and high-TB/HIV-burden countries, we propose that a spectrum of mechanisms underlies the development of TB during the first 3 months of ART. Much of this disease is likely to represent progression of subclinical TB that was present before ART initiation, either due to persisting immunodeficiency or due to unmasking during ART-induced immune recovery. The spectrum of effects of immune recovery on TB presentations in this period may include shortening of the time to development of symptoms (unmasking TB), increasing the rapidity of the initial onset of symptoms of TB, and increasing the intensity of clinical manifestations. These effects are likely to depend on the antigen load and the rate and intensity of immune recovery. Much remains to be learned from studies of the immunologic mechanisms underlying these phenomena.

Because it is presently not possible to clinically distinguish the underlying disease mechanisms, we suggest that cases of TB presenting during the initial months of immune recovery might collectively be referred to as “ART-associated TB.” To refer to
the phenomenon whereby immune recovery triggers the presentation of TB, we suggest the term “unmasking” disease be used. Finally, in agreement with the existing literature, we suggest that only the subset of patients with exaggerated and overtly inflammatory manifestations of TB should be referred to as IRD. To enable case definitions for unmasking TB and TB IRD to be derived, further immunologic and clinical studies are needed to better characterize these phenomena.

Conflict of Interest Statement: None of the authors has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

References

35. Lawn SD, Bekker LG, Miller RF. Immune reconstitution disease as a severe immune reconstitution disease following the commencement of highly active antiretroviral therapy. Sex Transm Infect 2003;79:337–338.


