Research paper

Controlling HIV among people who inject drugs in Eastern Europe and Central Asia: Insights from modelling

Peter Vickerman\textsuperscript{a,b,*}, Lucy Platt\textsuperscript{b}, Emma Jolley\textsuperscript{b}, Tim Rhodes\textsuperscript{b}, Michel D. Kazatchkine\textsuperscript{c}, Alisher Latypov\textsuperscript{d}

\textsuperscript{a} School of Social and Community Medicine, University of Bristol, UK
\textsuperscript{b} London School of Hygiene and Tropical Medicine, UK
\textsuperscript{c} UN Secretary General Envoy on HIV/AIDS in Eastern Europe and Central Asia, Geneva, Switzerland
\textsuperscript{d} Management Sciences for Health, Leadership, Management and Governance Project, 19 Maskovsky Prospect, 6th Floor, Room 605, Kiev, Ukraine

\textbf{A R T I C L E    I N F O}

\textbf{Article history:}
Received 20 June 2014
Received in revised form 24 September 2014
Accepted 26 September 2014

\textbf{Keywords:}
HIV
Eastern Europe
Central Asia
Modeling
PWID
NSP
OST
ARV
Combination interventions

\textbf{A B S T R A C T}

\textit{Background:} Although there is evidence of the effectiveness of needle and syringe programme (NSP), opioid substitution therapy (OST) and antiretroviral therapy (ART) in reducing HIV prevalence, most Central and Eastern European sub-regions still have low or no coverage of most or all of these interventions.

\textit{Methods:} We conducted a modelling analysis to consider the potential impact on HIV incidence and prevalence of OST, NSP and ART in three illustrative epidemic scenarios: Russia (St. Petersburg); Estonia (Tallinn) and Tajikistan (Dushanbe). For each intervention, we consider the coverage needed of each intervention separately or in combination to: (1) achieve a 30% or 50% relative reduction in HIV incidence or prevalence over 10 years; and (2) reduce HIV incidence to below 1% or HIV prevalence below 10% after 20 years. A sensitivity analysis for St. Petersburg considered the implications of greater or no risk heterogeneity, none or more sexual HIV transmission, like-with-like mixing, different injecting cessation rates and assuming a lower HIV acute phase cofactor.

\textit{Results:} For St. Petersburg, when OST, NSP and ART are combined, only 14% coverage of each intervention is required to achieve a 30% reduction in HIV incidence over 10 years. Similar findings are obtained for Tallinn and Dushanbe. In order to achieve the same reductions in HIV prevalence over 10 years, double the coverage level is required relative to what was needed to achieve the same reduction in HIV incidence in that setting. To either reduce HIV incidence to less than 1% or HIV prevalence to less than 10% over 20 years, with all interventions combined, projections suggest that very high coverage levels of 74–85% are generally required for the higher prevalence settings of Tallinn and St. Petersburg, whereas lower coverage levels (23–34%) are needed in Dushanbe. Coverage requirements are robust to increased sexual HIV transmission, risk heterogeneity and like-with-like mixing, as well as to assuming a lower HIV acute phase cofactor or different injecting cessation rate.

\textit{Conclusion:} The projections suggest that high but achievable coverage levels of NSP can result in large decreases (30%) in HIV incidence in settings with high HIV prevalence among PWID. Required coverage levels are much lower when interventions are combined or in lower prevalence settings. However, even when all three interventions are combined, the targets of reducing HIV incidence to less than 1% or prevalence to less than 10% in 20 years may be hard to achieve except in lower prevalence settings.

© 2014 Published by Elsevier B.V.

\textbf{Introduction}

Despite decreases in the rate of spread of HIV globally in the last decade, Eastern Europe and Central Asia have witnessed a dramatic increase in the incidence of HIV infection. Overall, the estimated number of adults and children living with HIV in this region in 2012 was 1.3 million [1,000,000–1,700,000] (UNAIDS, 2013). The epidemic is concentrated among populations at higher risk of HIV exposure, including people who inject drugs (PWID) and their sexual partners, men who have sex with men, sex workers, prisoners and migrants (Platt et al., 2013). In Eastern Europe and Central Asia, the shared use of drug injecting equipment has been the major driver of infections. Between 2006 and 2010, one in three

\textsuperscript{*} Corresponding author at: School of Social and Community Medicine, University of Bristol, UK.
E-mail address: peter.vickerman@bristol.ac.uk (P. Vickerman).

http://dx.doi.org/10.1016/j.drugpo.2014.09.013
0955-3959/© 2014 Published by Elsevier B.V.
of new HIV cases reported in Eastern Europe were associated with injecting drug use (Platt et al., 2013). The countries with the highest levels of reported diagnosed cases among PWID in Europe were Ukraine (153 per million people), Russia (98 per million people), and Kazakhstan (78 per million people) (Platt et al., 2013). Estimates from prevalence studies suggest that more than one in two PWID are HIV positive in parts of Estonia, Russia and Ukraine (Jolley et al., 2012; UNAIDS, 2013). Despite the recent increases in heterosexual transmission, the epidemic among PWID continues to expand, undefeated, and inadequately addressed in the region.

There is a large body of evidence demonstrating the effectiveness of combination interventions including needle and syringe programmes (NSP), opioid substitution therapy (OST) and antiretroviral therapy (ART) in reducing HIV incidence and prevalence among PWID (Abdul-Quader et al., 2013; Aspinall et al., 2014; MacArthur et al., 2012; Palmateer et al., 2010; van Den Berg, Smits, Van Brussel, Coutinho, & Prins, 2007; Wood et al., 2009). Indeed, a number of analyses from Central and Eastern Europe also suggest that these interventions could be cost-effective for decreasing HIV transmission (Alistar, Owens, & Brandleau, 2011; Kumaranayake et al., 2004; Long et al., 2006; Vickerman et al., 2006). Despite this, most Eastern European and Central Asian sub-regions of Eurasia still have low or no coverage of most or all of these interventions. OST is unavailable in Russia, Turkmenistan and Uzbekistan, and programmes in Estonia, Tajikistan and Ukraine are believed to reach only around 10%, 1% and 2% of PWID, respectively (Kurbatova, 2012; Latypov et al., 2012). The overall 35% coverage of ART in Eastern Europe and central Asia remains well below the global level of 60%. Furthermore, ART coverage is disproportionately low among PWID in Europe when compared with the general population (Donoghoe & Stengaard, 2010), and is particularly low in these high prevalence settings where the proportion of HIV positive PWID receiving ART is estimated to be much less than 10% (Mathers et al., 2010).

In such high HIV prevalence yet low intervention coverage settings, modelling can be useful for understanding what is required in terms of scaling up interventions to reduce HIV transmission to impact on the epidemic. Such model projections can be useful tools for influencing policy as has occurred for the debate on HIV and HCV treatment as prevention (Granich, Dye, Gilks, & De Cock, 2009; Grebely, Matthews, Lloyd, & Dore, 2013). To this end, we conducted a modelling analysis which considers the potential impact on HIV incidence and prevalence of OST, NSP and ART in three illustrative epidemic scenarios: Russia; Estonia and Tajikistan. The three interventions are key within the comprehensive package of interventions recommended by WHO, UNAIDS and UNODC for HIV prevention, treatment and care among PWID (WHO, 2012). Two of the epidemic scenarios are based on the high HIV prevalence (>40%) settings of St. Petersburg (Russia) and Tallinn (Estonia), whereas the third is based on a lower HIV prevalence (<20%) setting of Dushanbe (Tajikistan). All three settings currently have very low coverage of OST and ART among PWID at less than 10%. NSP coverage is high in Tallinn (approximately 70 syringes per PWID per year, Uuskula et al., 2011), moderate in Dushanbe (10–20 syringes per PWID per year (Personal communication, Ulugbek Aminov)) and low in St. Petersburg (Personal communication, Robert Heimer).

Methods

Overview of methods

We conducted a modelling analysis to consider the potential impact on HIV incidence and prevalence of OST, NSP and ART in three illustrative epidemic scenarios: Russia (St. Petersburg); Estonia (Tallinn) and Tajikistan (Dushanbe). The dynamic model of sexual and injection related HIV transmission among PWID is described in detail below. At baseline, the model is calibrated to detailed HIV prevalence and incidence data from each setting, adjusting for the possible decrease in HIV incidence resulting from heightened coverage of NSP in Tallinn (Uuskula et al., 2011) or moderate coverage of NSP in Dushanbe. The model also adjusts for possible longer duration of injecting in Tallinn and St. Petersburg than Dushanbe (Beyrer et al., 2009; Nicolai et al., 2010; Platt et al., 2006). In accordance with NSP data from Tallinn (Uuskula et al., 2011), the effect of NSP in Tallinn was assumed to scale up from 2003 to 2009 with the final efficacy estimated from fitting the model to observed prevalence and incidence trends in Tallinn, while assuming the efficacy in intermediate years is proportionate to the relative number of syringes distributed in that year compared to 2009. The same assumptions for the effect of NSP on HIV transmission were assumed for Dushanbe but with syringe distribution scaling up more slowly from nothing in 1999 to about 7 syringes per injecting drug user (IDU) per year in 2006, and then rapidly up to about 32 syringes per IDU per year by 2010 and 2011. The model was fit to HIV prevalence and incidence data for each setting by adjusting the HIV seeding prevalence in 1996 (to shift when the epidemic starts), the injecting related infection rate per month in the latent phase of HIV, and duration of injecting (both used to change the rate at which the epidemic progresses and the prevalence it stabilizes at). The sexual infection rate was also adjusted to give a prevalence of HIV among PWID in each setting that is due to sexual HIV transmission. The effect of NSP expansion in Tallinn was used to fit the model to the observed downturn in HIV incidence (and possibly prevalence) in Tallinn. The adjusted parameter values used for the model fits are presented in Table 1, while all other parameters were kept constant and are shown in Table 2. More details on the fitting methods are given below. A comparison of the model (Baseline projections) with prevalence and incidence data from each setting is shown in Fig. 1. It is important to note that the model runs should be seen as illustrative for the type of epidemic occurring in these different settings, i.e. the Tallinn epidemic represents a high prevalence epidemic with high coverage of NSP whereas the St. Petersburg and Dushanbe epidemics represent high and moderate HIV prevalence epidemics, respectively, with no or moderate NSP coverage.

Impact of scaling up OST, NSP and ART

These baseline model fits were then used to project the impact of scaling up OST, NSP and ART, while taking into account the following assumptions. Current receipt of OST was assumed to reduce an individual’s injection related probability of becoming infected by 50% based on a recent meta-analysis of cohort studies that estimates the reduction in HIV incidence among people currently on OST (MacArthur et al., 2012). Because PWID predominantly inject opiates in these settings (Abdala et al., 2008; Beyrer et al., 2009; Wilson, Sharma, Zilmer, Kalikova, & Uuskula, 2007), all PWID were assumed to be eligible for OST. Similarly, high coverage NSP (assumed to correspond to 70 syringes distributed per PWID per year as achieved in Tallinn in 2008/2009) was assumed to reduce an individual’s injection related risk of becoming infected by 40%, based on the possible effect of widespread NSP on HIV incidence in Tallinn (Uuskula et al., 2011) as calibrated through fitting the model to observed trends in HIV incidence in that setting. This effect is assumed to occur at the highest NSP coverage achieved in Tallinn in 2008/2009 (approximately 70 syringes distributed per IDU per year), whereas for lower coverage levels a linear relationship is assumed between syringe distribution per person per year and the relative decrease in transmission risk. This is likely to be a simplification of the real relationship between level of syringe distribution and resulting decrease in HIV incidence, but compares well with a recent meta-analysis of cohort studies that estimated exposure to NSP was associated with a 36–58% reduction in HIV
incidence among PWID (Aspinall et al., 2014). Any coverage of NSP is assumed to be relative to the maximum coverage of NSP achieved in Tallinn, with 100% coverage assumed to have the same efficacy as achieved in Tallinn in 2008/2009 (40% reduction in infection risk to all PWID) and 50% coverage assumed to have half this efficacy, i.e. 20% reduction in HIV infection risk among all PWID or 50% of PWID have a 40% reduction in risk. Any scale up of OST and NSP is assumed to occur over 7 years from 2012 to mimic the scale up of NSP in Tallinn (Uuskula et al., 2011), and the impacts of different final coverage levels are considered. Lastly, OST and NSP were assumed to not affect an individual’s sexual related probability of becoming infected.

Receipt of ART was assumed to reduce the sexual and injection related infectivity of an HIV positive PWID by 80%. This is based

Table 1
Behavioural and intervention coverage parameter estimates used in the model fits (symbols used in model equations in brackets).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Tallinn</th>
<th>St. Petersburg</th>
<th>Dushanbe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average duration inject in years (1/µ)</td>
<td>16</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>Injection related infection rate per month in latent phase of HIV (β_m)</td>
<td>0.014</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td>Sexual related infection rate per month in latent phase of HIV (β_w)</td>
<td>0.0023</td>
<td>0.0014</td>
<td>0.0014</td>
</tr>
<tr>
<td>Seed HIV prevalence in 1996 (y₀)</td>
<td>1.5%</td>
<td>0.1%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Percentage of PWID defined as high-risk</td>
<td>73%</td>
<td>5.4%</td>
<td>39%</td>
</tr>
<tr>
<td>Factor increase in injection related HIV transmission risk if high-risk</td>
<td>4.4</td>
<td>8</td>
<td>3.2</td>
</tr>
<tr>
<td>Baseline intervention coverage assumptions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSP (n(t))</td>
<td>Assumed to scale up from nothing in 2003 to 40% reduction in HIV incidence in 2008. Effect in intermediate years proportional to syringes distributed.</td>
<td>0%</td>
<td>Assumed to scale up from nothing in 1999 to about 20% reduction in HIV risk by 2010 because achieved half NSP coverage of Tallinn.</td>
</tr>
<tr>
<td>OST (o(t))</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>ART (recruitment rate r is varied)</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 2
HIV natural history and intervention efficacy parameters for model fits (symbols used in model equations in brackets).

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Value used</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV &quot;biological&quot; model parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injection related infection rate per month in latent phase of HIV (β_m)</td>
<td>Varied to fit model</td>
<td>See Table 1 for values used in model fit</td>
</tr>
<tr>
<td>Cofactor increase in HIV transmission probability during: Initial acute phase of high viraemia (δ)</td>
<td>26</td>
<td>Hollingsworth, Anderson, and Fraser (2008)</td>
</tr>
<tr>
<td>Pre-AIDS phase of high viraemia (θ)</td>
<td>7</td>
<td>Hollingsworth, Anderson, and Fraser (2008)</td>
</tr>
<tr>
<td>Duration of initial acute phase of high viraemia in years (1/δ)</td>
<td>0.25</td>
<td>Hollingsworth, Anderson, and Fraser (2008)</td>
</tr>
<tr>
<td>Duration of pre-AIDS phase of high viraemia in years (1/θ)</td>
<td>0.75</td>
<td>Hollingsworth, Anderson, and Fraser (2008)</td>
</tr>
<tr>
<td>Duration of latent phase in years (1/γ)</td>
<td>9.4</td>
<td>Prins et al. (1997)</td>
</tr>
<tr>
<td>Start time of HIV epidemic</td>
<td>1996</td>
<td>Start of HIV epidemic set to 1996 but seeding prevalence varied</td>
</tr>
<tr>
<td>Seed HIV prevalence at start of epidemic (y₀)</td>
<td>Varied to fit model</td>
<td>Estimated through fitting model to HIV prevalence and incidence data (Table 1)</td>
</tr>
<tr>
<td>Model intervention effectiveness parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative HIV infection rate while on ART compared to latent phase transmission probability (ω)</td>
<td>0.20</td>
<td>No data for PWID – estimated from recent trials and studies (Cohen et al., 2011; Donnell et al., 2010) adjusted for likely lower adherence levels among PWID (Bangsberg et al., 2000; Braithwaite et al., 2007; Gross et al., 2001; Malta et al., 2010; Nolan et al., 2011; Petersen et al., 2007; Wood et al., 2003) PWID have lower survival on ART than non-PWID (Carrico, 2011)</td>
</tr>
<tr>
<td>Average survival time with HAART in years (1/Δ)</td>
<td>15</td>
<td>MacArthur et al. (2012)</td>
</tr>
<tr>
<td>Relative infection rate if susceptible IDU is currently on OST (Ψ_r)</td>
<td>0.5</td>
<td>See text and (Aspinall et al., 2014; Uuskula et al., 2011)</td>
</tr>
<tr>
<td>High coverage NSP as in Tallinn in 2008/09 (Ψ_n)</td>
<td>0.6</td>
<td>Similar to recent study considering efficacy of OST and NSP combined for HCV (Turner et al., 2011)</td>
</tr>
<tr>
<td>OST and high NSP coverage (Ψ_m)</td>
<td>Product of above</td>
<td></td>
</tr>
</tbody>
</table>
on results of recent trials and prospective studies among serodiscordant heterosexual couples that have shown a 90% or greater reduction in HIV infectivity when one sexual partner is on ART (Cohen et al., 2011; Donnell et al., 2010), but adjusted downwards for the lower adherence levels frequently achieved among PWID (Malta, Magnanini, Strathdee, & Bastos, 2010; Nolan et al., 2011; Wood et al., 2003) which are likely to increase viral load (Bangsberg et al., 2000; Braithwaite et al., 2007; Gross, Bilker, Friedman, & Strom, 2001; Petersen et al., 2007). For simplicity, it is assumed that all HIV positive PWID (except those in the initial acute phase) can be recruited onto ART at a fixed rate. ART coverage is only measured among HIV positive PWID, and because the model’s projected impact of a specific proportion of PWID currently being on ART will be same irrespective of whether PWID transition on and off ART, for simplicity we assumed no loss to follow up. However, we do acknowledge that it does occur and will result in higher ART recruitment rates being needed to achieve the same overall ART coverage.

For each intervention, we considered the coverage needed of each intervention separately or in combination to firstly achieve a 30% or 50% relative reduction in HIV incidence or prevalence over 10 years; and secondly to reduce HIV incidence to below 1% or HIV prevalence below 10% after 20 years.

Technical model description

The model stratifies the population into those that are susceptible to HIV infection (stage $x$) and those that are HIV infected. The HIV infected population can either be in the initial high viraemia phase of infection (stage $h$ with average duration $1/\gamma$), a short late phase of high viraemia pre-AIDS (stage $a$ with average duration $1/\eta$), or on ART (stage $r$ with average duration $1/\Delta$). PWID enter the population at a rate $\Omega(t)$ that is set to maintain a constant population size before ART is initiated. PWID can be recruited onto ART (at a rate $r$) once they enter the long latent phase of HIV, upon which they have reduced infectivity (cofactor $o$). Those in the initial and late phases of high viraemia have heightened transmission (cofactors $\delta$ and $\theta$ respectively) compared to the injection and sexual related infection rate of those in the latent phase of HIV ($\beta_{inh}$ and $\beta_{sex}$). OST and NSP are assumed to have specific coverage levels ($n(t)$ and $d(t)$ that vary over time) and reduce injection related HIV transmission by cofactors $\psi_{nc}$ and $\psi_{ns}$ respectively, when not in combination, and by $\psi_{nc}$ if in combination. OST and NSP are not modeled explicitly because PWID move between these groups and so incorporating them as average coverage levels is a reasonable approximation. The model also stratifies the PWID into those with low and high injecting risk (denoted by the subscript $j=0$ for low risk and 1 for high risk, with $p_j$ being the initial proportion of PWID in each), with the injection related risk of HIV transmission among susceptible PWID in the high-risk strata being a factor ($m$) greater than among the low risk PWID. The model assumes a proportion ($z$) of the transmission events of PWID in a specific injecting risk state are with PWID from that same risk state (like-with-like mixing), and then the remaining transmission events are spread across PWID from any injecting risk state proportional to the overall relative frequency of transmission events for PWID in that state. Conversely, sexual HIV transmission among PWID is modelled simply with no heterogeneity and PWID mixing freely to form sexual contacts where sexual HIV transmission can occur. The model equations are included below:

\[
\begin{align*}
\frac{dx_0}{dt} &= \Omega(t)\mu_0 - (\Phi(t)\lambda_{inh}^0 + \lambda_{sex}k_0 - \mu)x_0 \\
\frac{dh_0}{dt} &= (\Phi(t)\lambda_{inh}^0 + \lambda_{sex}k_0 - h_0(v + \mu)) \\
\frac{dy_0}{dt} &= v h_0 - y_0(\mu + r) \\
\frac{da_0}{dt} &= 2r y_0 - a_0(\mu + r) \\
\frac{dt_0}{dt} &= r(a_0 + y_0) - t_0(\mu + \Delta) \\
\frac{dx_1}{dt} &= \Omega(t)p_1 - (\Phi(t)\lambda_{inh}^{-1} + \lambda_{sex}k_1 - \mu)x_1 \\
\frac{dh_1}{dt} &= (\Phi(t)\lambda_{inh}^{-1} + \lambda_{sex}k_1 - h_1(v + \mu)) \\
\frac{dy_1}{dt} &= v h_1 - y_1(\mu + r) \\
\frac{da_1}{dt} &= 2r y_1 - a_1(\mu + r) \\
\frac{dt_1}{dt} &= r(a_1 + y_1) - t_1(\mu + \Delta)
\end{align*}
\]
where $\Phi(t)$ is the overall protective effect of NSP and OST and has the following form (where the coverage of OST and NSP, $o$ and $n$, vary over time):

$$\Phi(t) = (1 - o - n + an) + o(1 - n)\psi_o + n(1 - o)\psi_n + an\psi_{ow}$$

And $\lambda_{sex}$ and $\lambda_{inj}$ are the sexual and injecting force of infection for HIV transmission which have the following form:

$$\lambda_{sex} = \frac{\beta_{sex}}{N} \sum_{i=0,1} \left( h_i \delta + y_i + \theta a_i + \omega r_i \right)$$

$$\lambda_{inj}^{0} = \frac{\beta_{inj}}{N_0 + mN_1} \left( \epsilon + (1 - \epsilon) \right) \left( mN_1 \right) \left( h_0 \delta + y_0 + \theta a_0 + \omega r_0 \right) / N_0$$

$$\lambda_{inj}^{1} = \frac{\beta_{inj}}{N_0 + mN_1} \left( 1 - \epsilon \right) \left( mN_1 \right) \left( h_1 \delta + y_1 + \theta a_1 + \omega r_1 \right) / N_1$$

where $N$ is the total PWID population size ($N=x+y+a+r$), $N_0$ and $N_1$ are the population sizes of the low and high risk groups, and $\epsilon$ is the degree to which PWID have injection related transmission events with PWID of the same risk strata. The inflow into the PWID population ($\Omega(t)$) is defined as below where $a$ is the number that would be in the AIDS state if no ART were present:

$$\Omega(t) = \mu N + \eta a$$

More details on model calibration

The model was first fit to data from St. Petersburg. The level of heterogeneity was parameterized using data from a prospective cohort study that found substantial heterogeneity in the incidence of HIV among PWID in St. Petersburg, with 5.4% of the PWID being assumed to be higher risk with an eight fold heightened transmission risk (Kozlov et al., 2006). This corresponded to the increased HIV transmission risk among PWID that frequently reported psychostimulant use in that study (Kozlov et al., 2006). The model was first fit to available HIV prevalence and incidence data assuming all HIV transmission in St. Petersburg is injection related by varying the injection related infection rate ($\beta_{inj}$), the seeding prevalence in 1996 and the duration of injecting ($1/\mu$). Then, because a large body of evidence suggest there could be considerable sexual HIV transmission occurring among PWID in St. Petersburg (Abdala et al., 2008; Kozlov et al., 2006; Nicolai, Shcherbakova, Toussova, Kozlov, & Heimer, 2009), we then calibrated the level of sexual HIV transmission in the model to produce the level of sexual HIV prevalence (7%) estimated to be occurring among PWID in St. Petersburg from a recent detailed modelling analysis (Mills, White, Colijn, Vickerman, & Heimer, 2013). The injecting related infection rate was then adjusted to re-calibrate the model to the HIV epidemiological data. This model calibration suggests that 14% of new HIV infections among PWID in St. Petersburg are due to sexual HIV transmission.

Unfortunately there is no comparable HIV incidence data that can be used to assess the determinants of HIV transmission for Tallinn or Dushanbe, nor is there any prior estimation of the proportion of HIV infections due to sexual HIV transmission in these settings. However, with respect to sexual risk, existing data suggests similar levels of sexual HIV transmission risk may be present in these settings, with a similar large proportion of PWID sexual partnerships being with PWID in St. Petersburg and Tallinn (about 50% in St. Petersburg. (Abdala et al., 2003a, 2003b; Nicolai et al., 2009; Shaboltas et al., 2006) versus 68% in Tallinn (Platt et al., 2006)), comparatively high frequencies of sexual partner change (median of 2 in last 6–12 months in Tallinn and Dushanbe, (Beyer et al., 2009; Platt et al., 2006; Wilson et al., 2007) compared to 1–3 in last 6 months in St. Petersburg (Nicolai et al., 2009; Shaboltas et al., 2006)), and low levels of condom use in all settings (60% never or occasionally use condoms in Tallinn, (Vorobjov et al., 2013) versus 57% of last sex acts being unprotected in St. Petersburg (Nicolai et al., 2009)). Lastly, the prevalence of HIV among potential sexual partners is high in both Tallinn and St. Petersburg, with 7.6% of FSWs being HIV infected in Tallinn in 2006 (Uuskula et al., 2010) and 5–15% of non-injecting sexual partners of PWID being infected in St. Petersburg in 2006 (Mills et al., 2013). Because of these similarities between the settings we assumed the same proportion of new HIV infections are due to sexual HIV transmission in all three settings. Similarly, because incorporating risk heterogeneity is likely to reduce the impact of each intervention, and studies from Tallinn and Dushanbe have shown there are a number of strong predictors of prevalent HIV infection (Beyrer et al., 2009; Platt et al., 2006; Uuskula et al., 2010), we decided to also incorporate some level of heterogeneity in injecting transmission risk in the Dushanbe and Tallinn model projections. For Tallinn, we reflected the increased risk associated with Fentanyl injection (4.4 times higher risk in 73% of the population, Platt et al., 2006; Uuskula et al., 2010), whereas in Dushanbe we reflected the increased risk associated with daily drug injecting (3.2 times higher risk in 39% of the population, Beyrer et al., 2009). We also assumed a low level of like-with-like mixing as found in other analyses of PWID (Mills et al., 2012).

The same model fitting method used in St. Petersburg was then used to fit the model to the HIV prevalence and incidence data available in Dushanbe and Tallinn. However, instead of fitting the risk of sexual HIV transmission to obtain a specific sexual HIV prevalence among the PWID, it was used to give 14% of new HIV infections being due to sexual HIV transmission as found in St. Petersburg.

Sensitivity analysis

A sensitivity analysis was undertaken to determine how the model impact projections may alter if we changed some of the important model structural and parameter assumptions. This analysis concentrated on the impact projections for St. Petersburg, and after refitting to the same HIV prevalence data considered how changes in model or parameter assumptions affected the projected required coverage of all three interventions for reducing HIV incidence below 1% or HIV prevalence below 10% over 20 years. The sensitivity analysis firstly considered the effect of fitting to the upper or lower bound HIV prevalence in 2006. It then considered the effect of assuming a shorter duration of injecting, 10 years instead of 30 years, in line with the current duration of injecting found in a cross-sectional study undertaken in St. Petersburg in 2006 (Nicolai et al., 2010). Then, because Kozlov et al. (2006) found substantial heterogeneity in the incidence of HIV among PWID in St. Petersburg, the effect of stratifying the PWID population into three injecting risk levels instead of two was considered, with a small subgroup (3%) of PWID having very high transmission risk (30 times baseline levels of transmission risk) and a larger subgroup (47%) having moderately high risk (6 times baseline risk levels). This corresponded to the increased HIV transmission risk among PWID associated with frequently reported psychostimulant use and having more than three sexual partners (Kozlov et al., 2006). This sensitivity analysis involved extending the model to include three levels of transmission risk in the same way that two levels were incorporated. Although no relevant data exists for St.
Petersburg, we also considered the effect of assuming greater assortative (like-with-like) mixing between PWID of different risk strata (50% instead of 20% of PWIDs form injecting partnerships with PWID of the same risk level, with the remainder being randomly distributed). This factor was considered because it has been shown to be important in previous analyses (Vickerman, Martin, Turner, & Hickman, 2012). Additionally, we considered the effect of incorporating no heterogeneity in transmission risk and assuming none or greater sexual HIV transmission – 17% HIV prevalence due to sexual HIV transmission instead of 7% (Mills et al., 2013). Lastly, because there is uncertainty in the HIV acute phase cofactor, and because it has been hypothesized to be important for determining the prevention impact of HIV treatment (Cohen et al., 2012; Kretzschmar, Schim van der Loeff, Birrell, De Angelis, & Coutinho, 2013), we also considered the effect of the model’s projections of a lower cofactor of 5 instead of 26 (Cohen et al., 2012; Hollingsworth, Anderson, & Fraser, 2008). In addition, to look into this effect further, we also projected the decrease in HIV incidence (after 20 years) resulting from achieving a 50% coverage of ART (among HIV positive PWID) for the model assuming a HIV acute phase cofactor of 5, 10, 15, 20 and 26, with the model being refit to HIV prevalence data from St. Petersburg for each scenario.

Results

Achieving a 30 or 50% relative decrease in HIV incidence or prevalence over 10 years

Fig. 2 shows the required coverage of different intervention combinations for achieving a 30 or 50% relative decrease in HIV incidence or prevalence among PWID compared to baseline over 10 years. Different combinations are considered for each setting because Tallinn already has high baseline coverage NSP, which is taken as the comparator for that setting, while Dushanbe has moderate coverage NSP. It is assumed that without any significant interventions, a setting such as St. Petersburg would first scale up NSP and then OST and/or ART. Because of this, we do not consider a scenario where OST or ART are scaled up in absence of NSP.

For St. Petersburg, the projections highlight that high coverage levels of NSP on its own (61% coverage for a 30% reduction in HIV incidence over 10 years) can achieve moderate (30%) decreases in HIV incidence over 10 years but much larger decreases (50%) are not possible – similar to what high NSP coverage has already achieved in Tallinn. However, if NSP is combined with ART or OST in St. Petersburg then larger decreases in HIV incidence are possible, with a 50% decrease in HIV incidence requiring 50–60% coverage of each intervention and a 30% decrease requiring a third to a half of the coverage needed with just NSP. When all three interventions are combined, the required coverage levels for each intervention reduce by a further 25–50%, with only 14% coverage of each intervention being required to achieve a 30% reduction in HIV incidence over 10 years in St. Petersburg.

Similar findings are obtained for Tallinn and Dushanbe except that greater impact is possible with lower coverage. For instance, the coverage required for a single additional intervention (OST or ART on top of the existing NSP) to reduce incidence by 30% is about a third of the levels required for a single intervention in St. Petersburg (just NSP on its own), and halving HIV incidence is now possible with a single additional intervention whereas this was not possible in St. Petersburg. This increased impact in Tallinn and Dushanbe can be explained by the shorter injecting duration as well as lower baseline HIV incidence in these settings in 2012 due both to the pre-existing moderate or high coverage NSP interventions in both settings and the lower overall risk in Dushanbe.

In order to achieve the same reductions in HIV prevalence over 10 years in any of the three settings, 1.5–2.6 times the coverage level is required relative to what was needed to achieve the same reduction in HIV incidence in that setting. This means high coverage levels of combined interventions are generally required to result in large reductions (50%) in HIV prevalence over 10 years.

Decreasing HIV incidence to less than 1% or HIV prevalence to less than 10% over 20 years

Fig. 3 considers the required coverage of each intervention combination to either reduce HIV incidence to less than 1% or HIV prevalence to less than 10% over 20 years. The projections for single or double interventions are not shown because they were generally not able to achieve these impact targets, except for Dushanbe where coverage levels of about 40–65% and 27–42% of any pair of interventions can reduce HIV incidence to <1% or prevalence to <10% after 20 years, respectively. However, when all three interventions are combined, these targets are now possible but still very high coverage levels are required (74–85% for each intervention) in the higher prevalence settings of Tallinn and St. Petersburg, although much lower coverage levels (23–34%) are possible in Dushanbe.

Sensitivity analysis

The sensitivity analysis in Fig. 4 shows the effect of changing specific model assumptions on the projected coverage level of each intervention required to reduce HIV prevalence to less than 10%
or HIV incidence to less than 1% in St. Petersburg over 20 years. The baseline model projections are also shown for comparison. The figure firstly shows that many of the changes to the model assumptions do not have much effect (<10% relative change) on the required intervention coverage levels, including incorporating: greater risk heterogeneity (3 strata of injecting risk instead of 2) or like-with-like mixing (50% instead of 20% of PWID mix with PWID of same injecting risk); greater sexual HIV transmission (17% instead of 7% HIV prevalence due to sexual HIV transmission among PWID); lower cofactor for acute HIV (5 instead of 26); reduced duration of injecting (10 years instead of 30 years); and fitting to the lower bound HIV prevalence in 2006 (39% instead of 44%). In contrast, incorporating no risk heterogeneity results in a large underestimation (by 20–30% in relative terms) of the required intervention coverage levels needed to achieve our impact targets, whereas not including sexual HIV transmission adds to this but to a lesser extent. Conversely, fitting the model to a higher HIV prevalence (55% instead of 44% in 2006) increases (by 15% in relative terms) our required intervention coverage levels. Lastly, although our sensitivity analysis showed little effect of assuming a lower acute phase cofactor on the combined intervention impact projections, Fig. 5 shows that the cofactor could still have a moderate effect on the preventative impact of HIV treatment on its own.

**Discussion**

The baseline model projections suggest that high but achievable coverage levels of NSP can result in large decreases (30%) in HIV incidence in settings with high HIV prevalence among PWID. Required coverage levels are much lower when interventions are combined or in lower prevalence settings, with only about 10%
coverage of NSP, OST and ART being required to achieve a 30% reduction in HIV incidence over 10 years in Dushanbe (similar findings for Tallinn and St. Petersburg). Even in the context of Tajikistan, with limited domestic HIV funding and trained human resources, there is no reason to believe that scaling up these HIV prevention and treatment responses to such a relatively low level of coverage would not be feasible.

The analysis also highlights the importance of combining interventions for reducing HIV incidence and prevalence to low levels in high prevalence settings, with no single or paired intervention (or only at high coverage in the lower prevalence setting of Dushanbe) being able to reduce HIV incidence to less than 1% or prevalence to less than 10% in 20 years. However, in combination these targets become possible, although still considerable, with about 80% of all three interventions being required to achieve these targets in Tallinn and St. Petersburg over 20 years, and about half of this (30–40% coverage of all three interventions) being required in Dushanbe. These high coverage targets are largely a product of the considerable heterogeneity in transmission risk existing among PWID populations in these settings, and points to the importance of targeting high-risk transmission groups for reducing HIV transmission to low levels. Interestingly, although sexual HIV transmission seems to be an important contributor to HIV transmission in these settings, our sensitivity analysis suggests that its presence does not substantially dampen the impact of OST and NSP in St. Petersburg because it only acts as an amplifier of existing high levels of injection related HIV transmission.

The legal and policy environment is critical to enabling the development of proven-to-be-effective combination HIV prevention interventions, including NSP, OST and ART, and our findings emphasize the critical importance of governments to support their implementation and scale-up. One of the strongest voices of policy resistance to OST emanates from Russia (Latykov et al., 2012; Rhodes, Sarang, Vickerman, & Hickman, 2010), where the use of methadone and buprenorphine in treating opioid dependence is legally prohibited. Resistance stems from efforts to preserve existing drug treatment systems alongside concerns to prevent the diversion of new medicines (such as methadone or buprenorphine) to the illicit market or safely monitor their use. More fundamentally, resistance to substitution treatments is grounded in the history of ‘narcology’, a subdivision of Soviet psychiatry, which conceives of treatment from addiction in abstinence terms (Latykov, 2011). Narcologists have opposed the use of methadone in opiate dependence treatment as a “toxic” substance creating severe addiction, and as one step removed from ‘legalising’ drug use (Elovich & Drucker, 2008; Rhodes et al., 2010). Conversely, resistance to NSP mostly stems from the fact that its implementation implicitly recognizes a population of injecting users that is illegal and criminalized in many countries in the region.

Laws and policies, as well as law enforcement practices, are critical aspects of the environment influencing HIV risk among PWID (Strathdee et al., 2010). In these contexts, structural interventions bringing about policy, legal or social change are required to enable sufficient HIV prevention scale-up. Ecological evidence indicates elevated HIV risk among PWID in settings without legal access to HIV prevention interventions such as OST and NSP compared to settings with access (Friedman et al., 2006; Hurley, Jolley, & Kaldor, 1997). The relaxation of legal restrictions to the provision of sterile needles and syringes is essential for increasing their availability and accessibility, reducing levels of risk behaviour, as well as potentially levels of police harassment among PWID (Booth et al., 2013; Cooper, Moore, Gruskin, & Krieger, 2005; Rhodes, Judd, et al., 2004; Wolfe, Carrieri, & Shepard, 2010). Despite this growing body of evidence demonstrating the potential HIV prevention impact of social and structural intervention approaches, the package of combination HIV prevention interventions promoted by WHO and other international agencies as core to national HIV prevention programming under-emphasises their role (Degenhardt et al., 2010; Gupta, Parkhurst, Ogden, Aggleton, & Mahal, 2008). HIV prevention approaches need to include interventions that foster legal and policy change to create an environment in which combination interventions are easily accessed.

The modelling described here has limitations, and so the projections should only be seen as indicative of the impact that could be expected from scaling up interventions in settings with different HIV prevalences. Firstly, the model only incorporates simply the effect of risk heterogeneity in relation to injecting. However, although the sensitivity analysis highlights the importance of incorporating risk heterogeneity it also suggests that incorporating greater heterogeneity may not be important for obtaining more realistic impact projections. Secondly, sexual HIV transmission among PWID is also modelled simply, but as highlighted by the sensitivity analysis this may not be an important limitation because its inclusion does not majorly affect the model projections even with heightened sexual HIV transmission. Thirdly, only single model fits were obtained for each setting – the sensitivity analysis suggests that allowing for parameter uncertainty and multiple model fits could better quantify the degree of uncertainty that is present in our impact projections, as also highlighted by previous analyses (Degenhardt et al., 2010; Strathdee et al., 2010), but are unlikely to affect our qualitative findings. Indeed, because the model is fit to multiple HIV prevalence and incidence estimates for each setting, the level of uncertainty due to not obtaining multiple model fits should be reduced as shown by the effect of fitting the model to the extreme bounds of the HIV prevalence estimates for St. Petersburg in 2006. Interestingly, although the effect of assuming a much lower HIV acute phase cofactor does not dramatically increase the impact of the combined interventions, it does affect the prevention impact of HIV treatment on its own. To better evaluate the impact of HIV treatment as prevention among high risk populations such as PWID, it is important to get a better handle on this parameter. Fourthly, minimal risk behavior data from each setting was used in the model fitting. This is because it is hard to parameterize a model using normal measures of syringe and equipment sharing because there is little data on the relative infectivity of these different modes of transmission and an unknown level of social desirability bias in their measures. Instead, the HIV prevalence and incidence data from each setting was used to calibrate the modelled HIV epidemic by adjusting the average monthly risk of injecting and sexual HIV transmission between any susceptible and infected IDU, the time at which the epidemic started, and the estimated leaving rate for HIV infected IDUs. The current duration of injecting reported in each setting was used to evaluate the likely difference in the injecting cessation rate for each setting while data from prior studies was used to estimate the likely degree of sexual HIV transmission and heterogeneity in HIV transmission risk. The model fits to the HIV prevalence and incidence data suggest that the model can portray the type of epidemic that occurred in each setting and the sensitivity analysis suggests that assuming different cessation rates or levels of sexual HIV transmission and risk heterogeneity is unlikely to change our results much.

Lastly, we did not consider uncertainty in the efficacy estimates for the different interventions. This is of concern for ART and NSP because there is firstly little evidence assessing the impact of ART on parenteral HIV transmission, and also it is hard to assess the efficacy of specific levels of syringe distribution on an individual’s risk of acquiring HIV. Despite this issue, the efficacy estimate for NSP seems reasonable because it coincides with the possible HIV impact of wide-spread NSP in Tallinn (Uuskula et al., 2011) and the impact of NSP on the transmission of HIV (Aspinall et al., 2014) and HCV (Nolan et al., 2014; Turner et al., 2011). It also seems
reasonable that ART will have a large impact on parenteral HIV transmission, as evidence shows a huge decrease in plasma viral load when individuals start treatment, and ecological studies have shown associations between PWID community viral load and HIV incidence at the population level in Vancouver and Baltimore (Kirk et al. 2011; Wood et al., 2009). Due to uncertainty in the exact effect of ART on HIV transmission in PWID, and because of the low adherence observed among PWID (Malta et al., 2010), we used a lower estimate of 80% for the efficacy of ART in reducing HIV transmission risk among PWID.

Two other modelling analyses have considered the impact of scaling up NSP, OST and ART among PWID (Degenhardt et al., 2010; Strathdee et al., 2010), with these suggesting that all three interventions need to be scaled up to high coverage (approximately 50% of HIV-infected PWID on ART and 50% of all PWID on OST or NSP in Degenhardt et al., 2010) to halve HIV incidence at or over five years. Although we assumed similar intervention efficacy estimates to these analyses, our impact projections are more optimistic, with approximately half the intervention coverage (about 25%) being required to halve HIV incidence after 10 years. The reason for these differences in impact projections is due to three factors. Firstly, we considered a longer time frame over which it is easier to achieve large decreases in HIV incidence, and secondly we evaluated the impact at the end of the time period instead of averaged over the whole intervention period as measured by Strathdee et al. (2010). If our analysis had considered 5 years instead of 10 years, then our St. Petersburg model projections suggest 35% coverage of each intervention would halve incidence at 5 years, whereas 61% coverage would halve the total number of incident HIV infections over 5 years. Lastly, the two previous analyses also assumed a significant proportion of PWID have no injecting transmission risk (up to half of PWID) because they do not share syringes, as well as a higher risk transmission group. This results in HIV transmission being overly concentrated among a small group of high-risk PWID within which it is harder for interventions to achieve impact, as also highlighted by our sensitivity analysis. However, although evidence definitely suggests there is heterogeneity in transmission risk among PWID in prospective cohort studies, there is little evidence to suggest that a large proportion of PWID have no injecting transmission risk even if they don’t report syringe sharing (Kozlov et al., 2006; Ruan et al., 2007; van Beek, Dwyer, Dore, Luo, & Kaldor, 1998; van Den Berg et al., 2007), and so the projections by Strathdee et al. (2010) and Degenhardt et al. (2010) may have underestimated intervention impact. Our study also adds to these previous analyses because firstly we ask the important question of what intervention combinations are needed to reduce HIV transmission to much lower levels in three specific Eastern European and Central Asian settings; and secondly we explicitly consider the implications of numerous important model assumptions to identify which are important and which are not.

Concluding remarks

In conclusion, although our analysis suggests that very high coverage levels of combined interventions are generally required to reduce HIV transmission among PWID to low levels over the next 20 years (HIV incidence to less than 1% or prevalence to less than 10%), moderate coverage levels of two or three interventions can still result in large impact (50% reduction in HIV incidence). To make this a reality, significant increases in the investment in prevention as well as policy changes are required throughout the Eastern European and Central Asian region, particularly of course in settings where OST remains illegal and where PWIDs hardly access ART. If this does not occur, then the HIV epidemic in these settings will continue to spread among PWIDs, their sexual partners and the general population, and many lives will be unnecessarily lost.

Acknowledgements

PV, TR and LP conceived the study. PV undertook all model analyses with extensive input from LP, EJ and AL. MK and TR input into implications of the model analyses. The manuscript was written by all authors. The initial modelling undertaken by PV was funded by the World Bank and an UK Medical Research Council New Investigators Award (G0716127). PV thanks the HIV Modelling Consortium for providing funding for his team.

Conflict of interest statement

No conflict of interest.

References


Synthesis of individual and neighbourhood-level factors. Sexually Transmitted Infections, 86(Suppl 3), iii79–iii84.


**Personal communication**

Personal communication, Robert Heimer.

Personal communication, Ulugbek Aminov.