Infectious Disease Modelling 3 (2018) 373–384

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

Use of routine HIV testing data for early detection of emerging HIV epidemics in high-risk subpopulations: A concept demonstration study

Houssein H. Ayoub [a, b, c], Susanne F. Awad [b], Laith J. Abu-Raddad [b, c, d, *]

[a] Department of Mathematics, Statistics, and Physics, Qatar University, Doha, Qatar
[b] Infectious Disease Epidemiology Group, Weill Cornell Medicine-Qatar, Cornell University, Qatar Foundation - Education City, Doha, Qatar
[c] Department of Healthcare Policy and Research, Weill Cornell Medicine, Cornell University, New York, USA
[d] College of Health and Life Sciences, Hamad Bin Khalifa University, Doha, Qatar

ARTICLE INFO

Article history:
Received 22 March 2018
Received in revised form 11 June 2018
Accepted 18 October 2018
Available online 6 November 2018
Handling Editor: James Koopman

Keywords:
HIV
High-risk subpopulation
Surveillance
Mathematical modeling
Monte Carlo simulations
Epidemiology
Sexually transmitted infection

ABSTRACT

Introduction: HIV epidemics in hard-to-reach high-risk subpopulations are often discovered years after epidemic emergence in settings with poor surveillance infrastructure. Using hypothesis-generation modeling, we aimed to investigate and demonstrate the concept of using routine HIV testing data to identify and characterize hidden epidemics in high-risk subpopulations. We also compared this approach to surveillance based on AIDS case notifications.

Methods: A deterministic mathematical model was developed to simulate an emerging HIV epidemic in a high-risk subpopulation. A stochastic Monte Carlo simulation was implemented on the total population to simulate the sampling process of generating routine HIV testing data. Epidemiological measures were estimated on the simulated epidemic and on the generated testing sample. Sensitivity analyses were conducted on the results.

Results: In the simulated epidemic, HIV prevalence saturated at 32% in the high-risk subpopulation and at 0.33% in the total population. The epidemic started its emerging-epidemic phase 28 years after infection introduction, and saturated 67 years after infection introduction. In the simulated HIV testing sample, a significant time trend in prevalence was identified, and the generated metrics of epidemic emergence and saturation were similar to those of the simulated epidemic. The epidemic was identified 4.0 (95% CI 3.4–4.6) years after epidemic emergence using routine HIV testing, but 29.7 (95% CI 15.8–52.1) years after emergence using AIDS case notifications. In the sensitivity analyses, none of the sampling biases affected the conclusion of an emerging epidemic, but some affected the estimated epidemic growth rate.

Conclusions: Routine HIV testing data provides a tool to identify and characterize hidden and emerging epidemics in high-risk subpopulations. This approach can be specially useful in resource-limited settings, and can be applied alone, or along with other complementary data, to provide a meaningful characterization of emerging but hidden epidemics.

* Corresponding author. Weill Cornell Medical College – Qatar, Qatar Foundation - Education City, P.O. Box 24144, Doha, Qatar.
E-mail address: lja2002@qatar-med.cornell.edu (L.J. Abu-Raddad).
Peer review under responsibility of KeAi Communications Co., Ltd.

https://doi.org/10.1016/j.idm.2018.10.001

1. Introduction

HIV epidemics emerge first in specific subpopulations whose risk-behavior characteristics expose them to a higher risk of HIV infection (Anderson & May 1991). In most countries, HIV epidemics remain concentrated in these high-risk subpopulations including men who have sex with men (MSM), people who inject drugs (PWID), and female sex workers (FSWs) and their clients (Abu-Raddad et al., 2010a; May, Anderson, & Irwin, 1988; Walker et al., 2004). Due to the generally hidden nature, social stigma, and legal persecution of these subpopulations, it is difficult to capture emerging HIV epidemics among them by conventional and passive surveillance systems (Mills et al., 2004). Concentrated HIV epidemics in these high-risk groups are often detected years after epidemic emergence (Mumtaz et al., 2011, 2014a, 2014b). The delay in detecting epidemics as they emerge increases the risk of larger HIV epidemics emerging, and may limit access to HIV prevention services when needed (Mills et al., 2004).

The Middle East and North Africa (MENA), for instance, is a region characterized by recently emerging HIV epidemics among MSM and PWID (Mumtaz et al., 2011, 2014a, 2014b), and to a lesser extent among FSWs and clients (Abu-Raddad et al., 2010a, 2010b). Most of HIV incidence appears to be occurring among these populations or their direct contacts (such as spouses of PWID or spouses of clients of FSWs) (Kouyoumjian et al., 2018; Mumtaz et al., 2013, 2018). Most epidemics in this region, extending from Morocco in the west to Pakistan and Afghanistan in the east, were discovered years after they emerged (Mumtaz et al., 2011, 2014a, 2014b). In Libya, for example, anecdotal evidence suggested the possibility of a large HIV epidemic among PWID since around 2000 (Abu-Raddad et al., 2010a, 2010b). Nevertheless, the stigma and marginalization of PWID prevented direct access to this population till recently, when it was discovered that Libya is enduring one of the largest HIV epidemics ever among PWID, at a prevalence of 87% (Mirzoyan et al., 2013).

Novel surveillance methodologies have been developed to study HIV epidemiology among hidden and hard-to-reach populations (Berchenko & Frost, 2011; Diaz et al., 2009; Heckathorn, 1997; Mills et al., 2004; Rehle et al., 2004). Repeated integrated bio-behavioral surveillance surveys incorporating state of the art sampling methodologies for reaching hidden populations, such as respondent driven sampling, provide the best approach to study these key subpopulations (Magnani et al., 2005; Joint United Nations Programme on HIV/AIDS, 2000). However, the commitment of national programs to working with such politically-sensitive populations is a persistent challenge (Abu-Raddad et al., 2010a, 2013; Bozicevic, Riedner, & Calleja, 2013; Mumtaz et al., 2011). These methodologies can also be costly to implement in resource-limited settings, and the technical expertise to conduct them may not be available. Even when implemented, in most cases they are executed infrequently preventing longitudinal inferences about the epidemics, and often applied only in select geographic areas (Mumtaz et al., 2011, 2014a).

Against this background, we attempt here to answer the following question: Is it possible to identify emerging HIV epidemics among hidden and hard-to-reach high-risk subpopulations using routine and readily available HIV testing data such as those of blood donors among others (which are typically not used for surveillance purposes)? To address this question, we conducted hypothesis-generation modeling simulations to demonstrate this concept for HIV surveillance (study concept is illustrated in Fig. 1), thereby offering potentially a rapid and inexpensive methodology for detecting HIV epidemics as they emerge. If there are other supplementary data available, this approach may also be helpful to provide rough estimates for the epidemic growth rate and population size of the high-risk group, where the epidemic is expanding.

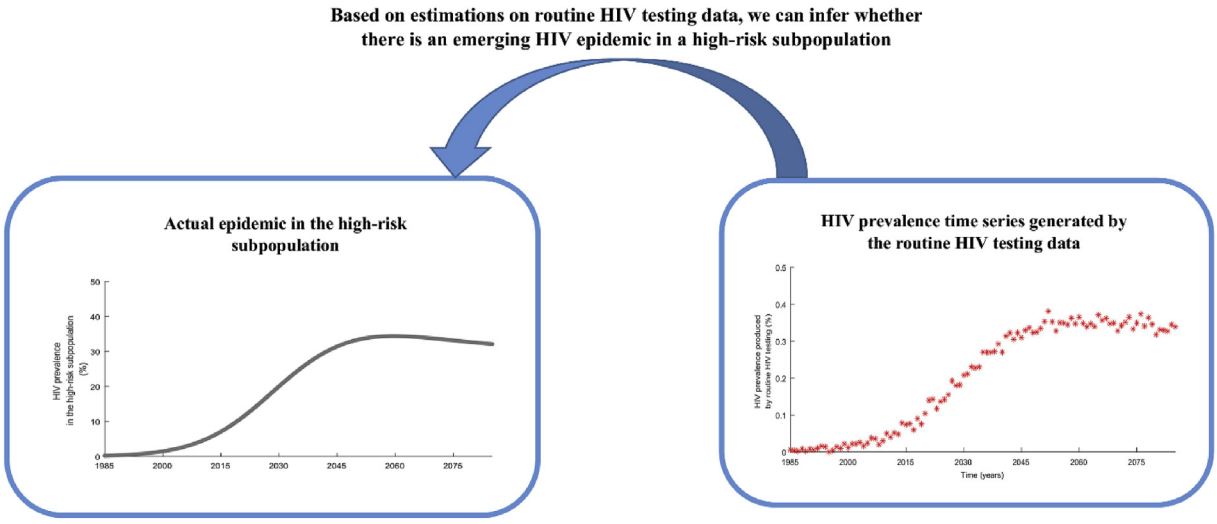

Fig. 1. Concept of the study. A schematic diagram of the use of routine HIV testing data to identify an emerging HIV epidemic in a high-risk subpopulation.

The utility of this approach hinges upon the availability of routine HIV testing data. While HIV testing of blood donors is available universally, and for different types of blood donation, different countries may have different protocols, and some may have stringent criteria for accepting blood donations that may exclude high-risk populations (Suligoi et al., 2010). While we use blood donor data as an example of routine testing data, our approach can generically be applied for any type of routine HIV testing data.

Indeed, a main strength of the present study is the availability of a large volume of HIV testing data in MENA (Abu-Raddad et al., 2010a, 2010b; Hermez et al., 2010), the regional focus of our study. Though methodological HIV surveillance continues to be rather limited in this region (Bozicevic et al., 2013), there is (probably often unnecessarily) high emphasis on HIV testing of broad general populations (Abu-Raddad et al., 2010a; Hermez et al., 2010). Reviews of current practices have indicated routine testing in diverse populations such as pregnant women, marriage applicants, university students, public sector employees, out-migrant (for visa to work abroad), in-migrants (for residency or visa renewal), prisoners, tuberculosis patients, and sexually transmitted disease clinic attendees (Abu-Raddad et al., 2010a; Hermez et al., 2010). For instance, more than 53 million HIV tests have been conducted in MENA within a short time horizon (Abu-Raddad et al., 2010a). As a specific country example, >500,000 tests are routinely conducted in Qatar every year (Ministry of Public Health, 2017), a country of less than three million inhabitants (Ministry of Development Planning and Statistics, 2015). This volume of testing data in MENA is yet to be utilized for HIV surveillance or scientific analysis purposes.

2. Methods

2.1. Mathematical model of an emerging HIV epidemic in a high-risk subpopulation

2.1.1. Mathematical model structure

A deterministic dynamical mathematical model, based on extension and adaptation of earlier models (Abu-Raddad, Patnaik, & Kublin, 2006; Abu-Raddad & Longini, 2008), was constructed to describe HIV transmission in a given population. The model consisted of a system of coupled nonlinear differential equations that stratified the population according to HIV status, stage of infection, and risk group. HIV progression was represented by three stages: acute, latent, and advanced. The model was parameterized to represent an emerging epidemic in a hidden high-risk subpopulation such as MSM, but it can be generalized to represent other hidden subpopulations such as FSWs and clients and PWID.

The model describes an HIV epidemic in a population of 100,000 individuals, representative of a population of a metropolitan area. The population was divided into 20 risk groups, from lower to higher levels of risk behavior. It was assumed that most individuals belong to the lower risk groups (general populations), and the size of the high-risk subpopulation where the HIV epidemic is emerging (say MSM) did not exceed 1% of the total population.

To describe the mixing between different risk groups, we included a mixing matrix that incorporated both an assortative component (choosing partners only from within the same risk group), and a proportionate component (choosing partners with no preferential bias based on the risk group). HIV was introduced in the population in 1985, and the model simulated the spread of the infection over a period of 100 years.

Further details about the model structure can be found in the Supplementary Material (SM).

2.1.2. Model parameterization

The model was parameterized using quality natural history and epidemiological data for HIV. For example, the durations of the acute, latent, and advanced stages were assumed to be 49 days (acute), 9 years (latent), and 2 years (advanced). These choices were based on compilation of data by the Joint United Nations Programme on HIV/AIDS (UNAIDS) (UNAIDS; UNAIDS/WHO, 2010a,b), and based on the classification in Wawer et al. (Wawer et al., 2005), re-analysis of the Rakai data for acute infection (Pinkerton, 2008), and measured time from seroconversion to death in cohort studies (UNAIDS/WHO, 2010a,b). The model's parameter values along with their references are listed in Table S1 in SM. The risk of exposure to HIV infection in each risk group was parameterized by the effective partnership change rate of each risk group (Awad & Abu-Raddad, 2014; Awad et al., 2015).

To simulate an emerging epidemic in the high-risk subpopulation that is *not sustainable* in the general population, the per-risk group effective partnership change rate was set to be large enough to generate a sustainable epidemic only in the four highest risk groups (risk groups 17–20). These risk groups represent different levels of risk behavior within the high-risk subpopulation. The remaining risk groups (risk groups 1–16) were considered as lower risk groups (general populations) in which the effective partnership change rate was not large enough to sustain an epidemic without mixing with the core high-risk subpopulation.

Further details about parameter assumptions can be found in SM.

2.1.3. Characterizing the actual epidemic using epidemiological summary measures

We estimated key epidemiological measures from the actual epidemic using the mathematical model such as HIV prevalence in the total population and HIV prevalence in the high-risk subpopulation. The trend of an epidemic is rather complex; therefore, to characterize the actual epidemic in terms of key epidemiological measures such as the time of

epidemic emergence, epidemic growth rate, and time of epidemic saturation, we implemented a simple approach by fitting the actual epidemic to the following segmented linear model (Vickerman et al., 2010):

$$\text{HIV prevalence at time } t = \begin{cases} c & t \le a \\ \beta(t-a)+c & a < t \le b \\ d & t > b \end{cases} \quad (1)$$

here, c and d are constants, the breakpoint a is the time of epidemic emergence, β is the epidemic growth rate, and the breakpoint b is the time of epidemic saturation.

2.1.4. Additional epidemiological measures

Assuming the availability of other supplementary data, the approach proposed here can potentially provide rough estimates regarding additional epidemiological measures such as epidemic growth rate and population size of the high-risk subpopulation where the epidemic is emerging. To generate these estimations, we used an analytical approximation to link the epidemic growth rate in the total population with the epidemic growth rate *only* in the high-risk subpopulation. This link is expressed by the following equation:

$$\beta_T \approx frac_{HR}\beta_{HR} \quad (2)$$

where β_T is the epidemic growth rate in the total population, β_{HR} is the epidemic growth rate in the high-risk subpopulation, and $frac_{HR}$ is the fraction of the population that belongs to the high-risk subpopulation. Derivation of this equation can be found in Section 3 of SM.

From this mathematical expression it is possible to roughly estimate the epidemic growth rate in the high-risk subpopulation, *if* the fraction of the population that belongs to the high-risk subpopulation is known. Likewise, it is possible to roughly estimate the population size of the high-risk subpopulation, *if* the epidemic growth rate in the high-risk subpopulation is known. Of notice that risk group size estimates can often be available, despite the limited availability of HIV data, as these data could have been generated for other purposes, such as the case for PWID whose estimates are often generated for drug-related purposes (Mumtaz et al., 2014a). Moreover, even when such data are not available, regional or global estimates can be used as informed rough estimates of the size of the subpopulation, such as the case for MSM (Caceres et al., 2008; McFarland & Caceres, 2001; Mercer et al., 2009; Mumtaz et al., 2011; UNAIDS).

2.2. Simulation of routine HIV testing data and AIDS case notifications

2.2.1. Stochastic Monte Carlo simulation of routine HIV testing and AIDS case notifications

We constructed a stochastic Monte Carlo simulation to simulate routine HIV testing data, such as those performed in different countries in populations including blood and organ donors, pregnant women, and marriage applicants among others (Abu-Raddad et al., 2010a; Hermez et al., 2010). The simulation consisted of randomly sampling every year 2% (2000 individuals) of the total population, which is a good approximation of the percentage of individuals who donate blood per year (World Health Organization, 2012). For example, in high income countries such as those in North America and Europe, the average percentage of the population that donate blood is 3%; whereas in some MENA countries the average is between 1% and 2% (World Health Organization, 2012).

In the sampling process, it was assumed initially that each individual has the same probability of being sampled, independent of their risk group. Subsequently, as part of sensitivity analyses, we studied the effects of biases in the sampling process.

The sampling process was repeated 1000 times with the results reported as summary measures on these 1000 realizations.

2.2.2. Characterizing the sampled population using epidemiological summary measures

HIV prevalence and its 95% confidence interval (CI) were estimated for each sampled year and compared with the actual epidemic. The time of epidemic emergence and saturation as well as epidemic growth rate were also estimated from the sampled population by fitting the segmented linear model described in Equation (1) to HIV prevalence time series.

We conducted, every year, a chi-square trend test on HIV prevalence time series up to that year, as estimated on the sampled population, to identify the temporal trend of HIV prevalence. We also conducted, every year, a chi-square trend test on AIDS case notifications time series up to that year, as also estimated on the sampled population, to identify the temporal trend in AIDS case notifications.

We recorded, for each statistical analysis, the earliest year in which the trend in HIV prevalence was found to be increasing with statistical significance. We also recorded, for each statistical analysis, the earliest year in which the trend in AIDS case notifications was found to be increasing with statistical significance. The latter represents passive surveillance of AIDS cases, as was common in the 1980s and 1990s, as a means to identify an emerging epidemic. This was done for merely comparison

purposes, to highlight the differences with routine HIV testing surveillance. Of notice that AIDS surveillance refers here strictly to our ability to detect a statistically-significant trend of an emerging HIV epidemic using only AIDS cases.

2.2.3. Sensitivity analyses

Sensitivity analyses were conducted to evaluate the impact of variation in the routine HIV testing sample size (and AIDS case notifications), as well as biases in the sampling process, on the representativeness and accuracy of the epidemiological inferences based on the sampled population.

The biases in routine testing data can arise for different reasons, such as different testing protocols for different populations or regions, that may effectively underrepresent or overrepresent specific high-risk subpopulations (Suligoi et al., 2010). The biases can also arise because of inherent differences in the underlying population, such as the differences between different blood donors—paid versus family replacement versus voluntary non-remunerated donors (Abdel Messih et al., 2014; Suligoi et al., 2010).

We conducted simulations with a smaller HIV testing sample size (1% of the total population instead of 2%), and larger sample size (5% of the total population instead of 2%). The effect of undersampling or oversampling the high-risk subpopulation was assessed by including a probability function in the random sampling process that underweighted, or overweighted the probability of randomly choosing an individual from the high-risk subpopulation.

3. Results

3.1. Actual epidemic emerging in the high-risk subpopulation

3.1.1. Epidemiological summary measures from the actual epidemic

HIV infection was introduced in the year 1985 in the high-risk subpopulation. The epidemic reached saturation at an HIV prevalence of 32% in the high-risk subpopulation (Fig. 2B), and at a prevalence of 0.33% in the total population (Fig. 2A). HIV prevalence in the four highest risk groups within the high-risk subpopulation ranged from 25% (risk group 17) to 48% (risk group 20) (Fig. 2C). The fraction of the high-risk subpopulation out of the total population (risk groups 17–20 combined) was 0.97% (Fig. 2D).

We estimated from the segmented linear model fitting of the actual epidemic (using Equation (1), Fig. 2A) that the epidemic started its rapid epidemic phase 28 (95% CI 27–29) years after the introduction of the infection, and saturated 67 (95% CI 66–68) years after the introduction of the infection (Fig. 2A). The estimated epidemic growth rate of the actual epidemic in the total population was 8.4×10^{-3}(95% CI $8.2 \times 10^{-3} - 8.7 \times 10^{-3}$) per year, and in the high-risk subpopulation it was 0.84(95% CI 0.82 – 0.87) per year—that is 100 times that in the total population.

3.1.2. Additional epidemiological measures

Assuming that the population size of the high-risk subpopulation is known from some other supplementary data source (here by design it was set in the model at about 1% of the total population), we can infer using Equation (2) that the growth rate in this high-risk subpopulation β_{HR} is equal to 0.84 per year. Alternatively, assuming that the growth rate in the high-risk subpopulation (β_{HR}) is what is known from some other supplementary data, then we can infer using Equation (2) that the population size of the high-risk subpopulation is about 1% of the total population.

3.2. Simulation of routine HIV testing data and AIDS case notifications

3.2.1. Epidemiological summary measures from the sampled population

Fig. 3 shows HIV prevalence and its corresponding 95% CI in the population sample generated by the simulated stochastic HIV routine testing on the total population (such as blood donation). Fig. 3A demonstrates that the sample-estimated HIV prevalence was representative of HIV prevalence in the actual epidemic. For the year 2013, when the epidemic started its rapid epidemic phase, HIV prevalence in the sample was 0.031% (95% CI 0.013%–0.079%), whereas HIV prevalence in the actual epidemic was 0.039%. Sometime during the rapid growing phase of the epidemic, for example in the year 2032, HIV prevalence in the sample was 0.16% (95% CI 0.11%–0.24%), whereas HIV prevalence in the actual epidemic was 0.18%. For the year 2052, the year when the epidemic saturated, HIV prevalence in the sample was 0.32% (95% CI 0.26%–0.41%), whereas HIV prevalence in the actual epidemic was 0.33%.

The segmented linear model fitting for the population sample is illustrated in Fig. 3B. The time after infection introduction but before epidemic emergence was estimated at 28 years (95% CI 25–32) (Table 1)—the same year [28 (95% CI 27–29)] as estimated from fitting the actual epidemic (Fig. 2A and Table 1). Epidemic saturation was reached 66 years (95% CI 62–70) after infection introduction—close to the year [67 (95% CI 66–68)] as estimated from fitting the actual epidemic (Fig. 2A and Table 1). The epidemic growth rate in the total population was 8.4×10^{-3}(95% CI $7.2 \times 10^{-3} - 9.9 \times 10^{-3}$) per year—similar to the epidemic growth rate estimated from fitting the actual epidemic [8.4×10^{-3}(95% CI $8.2 \times 10^{-3} - 8.7 \times 10^{-3}$) per year] (Fig. 2A and Table 1).

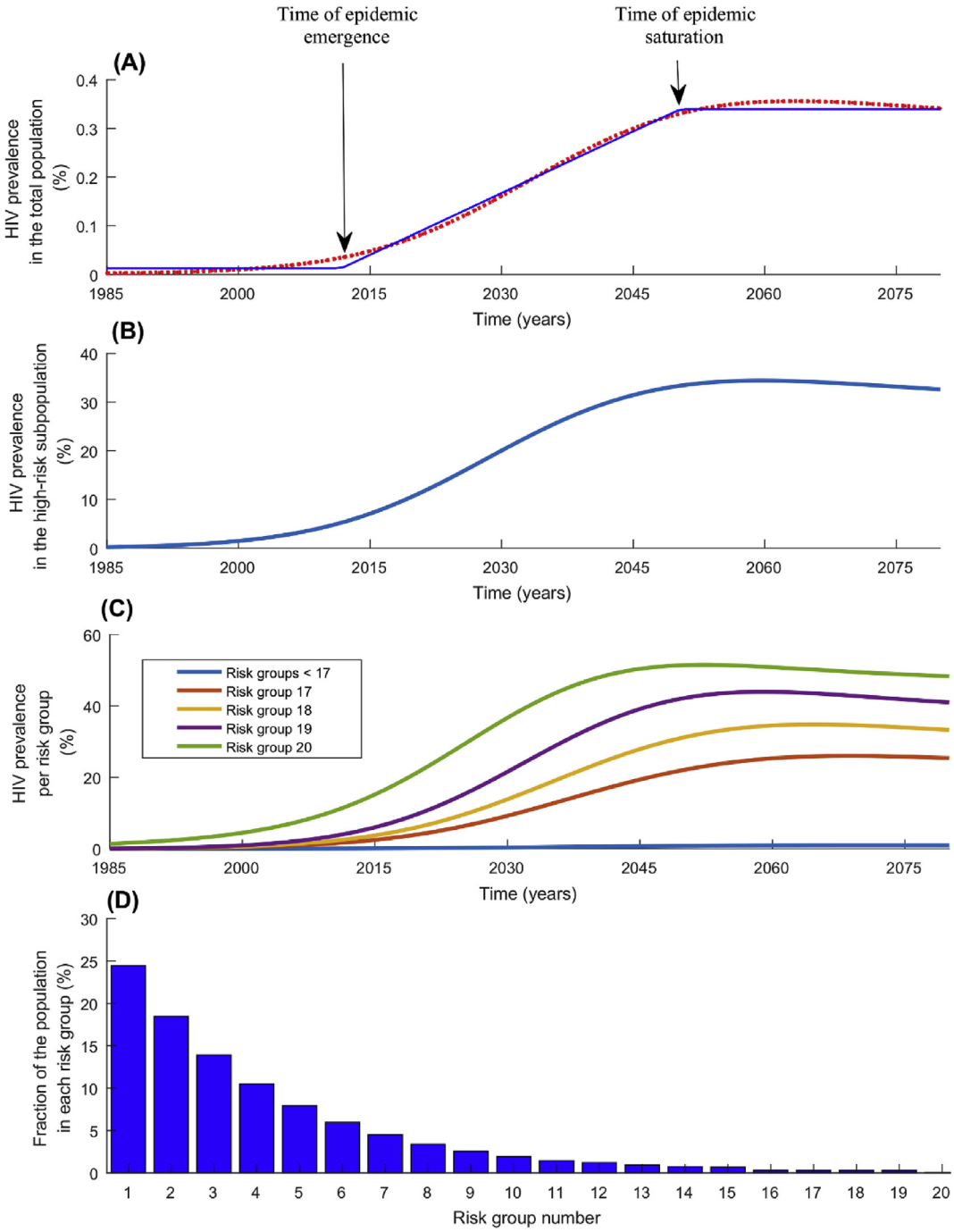

Fig. 2. Epidemiological summary measures for the actual HIV epidemic. A) HIV prevalence in the total population along with the segmented linear model fit. B) HIV prevalence in the high-risk subpopulation. C) HIV prevalence in the different risk groups in the population. D) Proportion of the population in each risk group.

Fig. 4 shows the year of epidemic identification (after epidemic emergence) using the HIV prevalence times series generated using the routine HIV testing, versus the year of epidemic identification using the time series of AIDS case notifications. Using routine HIV testing, the emerging epidemic was identified (with statistical significance) 4.0 (95% CI 3.4–4.6) years after epidemic emergence. Meanwhile, using AIDS case notifications, the emerging epidemic was identified (with statistical significance) 29.7 (95% CI 15.8–52.1) years after epidemic emergence. Across the 1000 realizations, routine HIV testing identified the emerging epidemic 24.3 (95% CI: 11.9–49.8) years before AIDS case notifications.

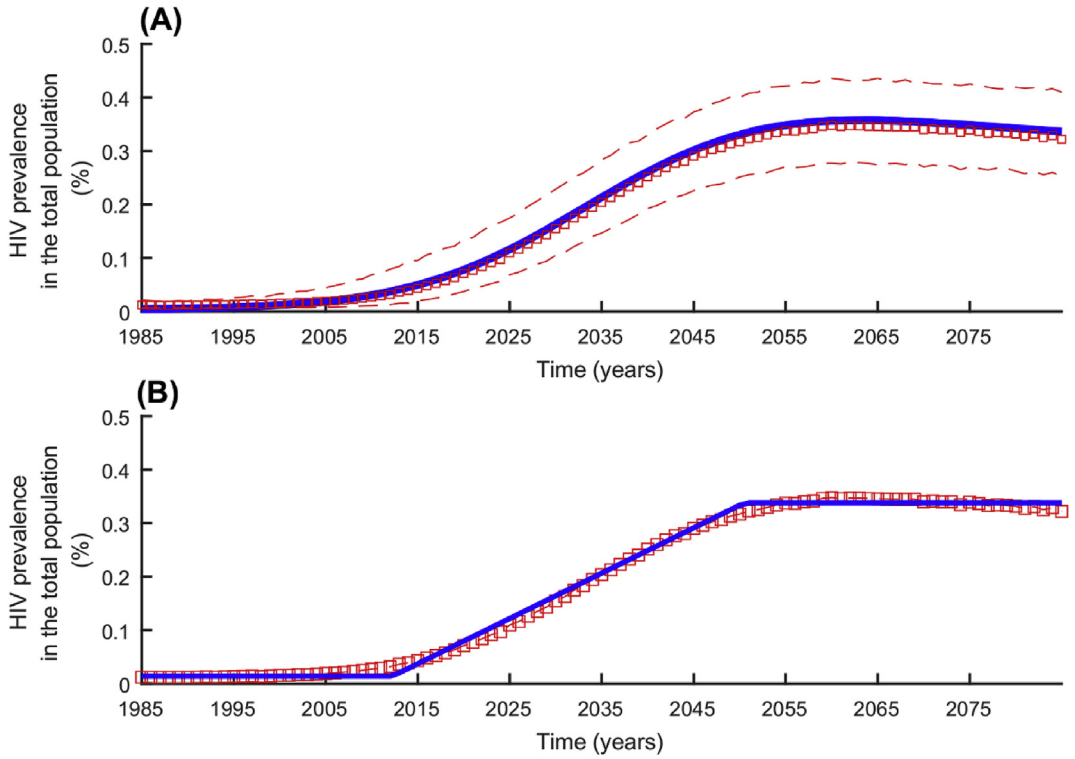

Fig. 3. Epidemiological measures for the simulated routine HIV testing sampled population. A) Mean and 95% confidence interval of the HIV prevalence time series estimated from the sampled population (red boxes), compared to HIV prevalence in the actual epidemic (blue line). B) The segmented linear model fit of HIV prevalence time series as estimated from the sampled population (red boxes). All of these results are for the total population.

Table 1
Estimated measures using the segmented linear model fitting of the actual epidemic and the simulated routine HIV testing sampled population. The Table includes also the results of the sensitivity analyses.

	Number of years to epidemic emergence after HIV introduction (95% CI)	Number of years to epidemic saturation after HIV introduction (95% CI)	Epidemic growth rate (95% CI)
Actual epidemic in the total population	28 (27–29)	67 (66–68)	$8.4 \times 10^{-3} (8.2 \times 10^{-3} - 8.7 \times 10^{-3})$
Simulated routine HIV testing sampled population (2% of the total population)	28 (25–32)	66 (62–70)	$8.4 \times 10^{-3} (7.2 \times 10^{-3} - 9.9 \times 10^{-3})$
Sensitivity analysis of smaller sample size for the simulated routine HIV testing sampled population (1% of the total population)	29 (24–34)	68 (63–72)	$8.6 \times 10^{-3} (6.8 \times 10^{-3} - 10.7 \times 10^{-3})$
Sensitivity analysis of larger sample size for the simulated routine HIV testing sampled population (5% of the total population)	28 (26–30)	67 (66–69)	$8.5 \times 10^{-3} (7.7 \times 10^{-3} - 9.3 \times 10^{-3})$
Sensitivity analysis of undersampling the high-risk subpopulation in the simulated routine HIV testing sampled population	29 (25–35)	66 (61–72)	$4.5 \times 10^{-3} (3.5 \times 10^{-3} - 5.6 \times 10^{-3})$
Sensitivity analysis of oversampling the high-risk subpopulation in the simulated routine HIV testing sampled population	28 (26–30)	67 (65–70)	$1.7 \times 10^{-2} (1.5 \times 10^{-2} - 1.9 \times 10^{-2})$

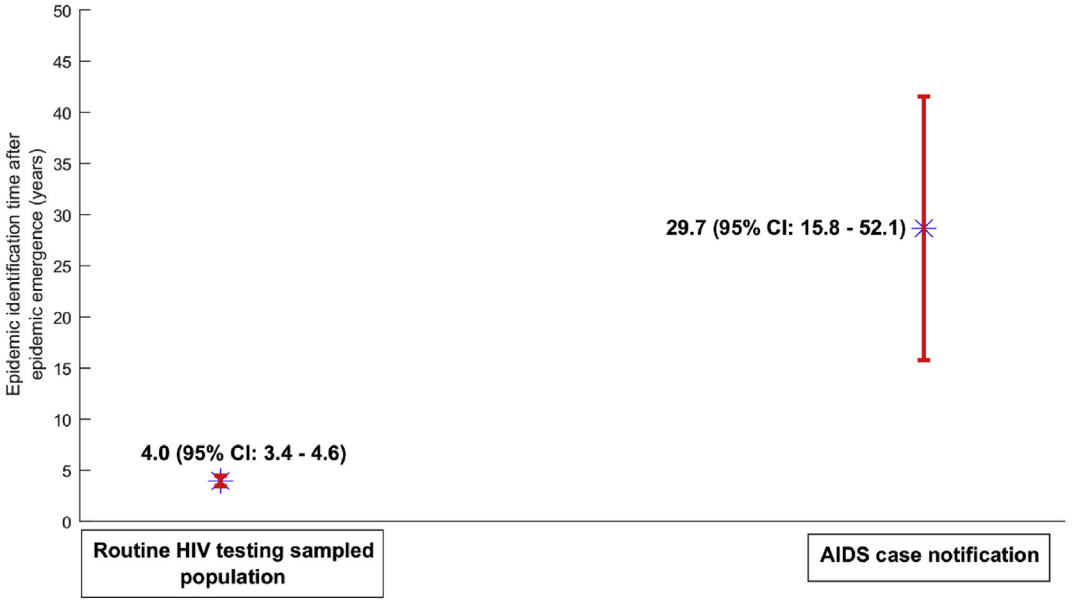

Fig. 4. Epidemic identification time. Mean and 95% confidence interval of the year of detecting the epidemic after epidemic emergence using the routine HIV testing sampled population versus using AIDS cases notification.

3.2.2. Sensitivity analyses

None of the sampling biases included in this study affected the conclusion of an emerging epidemic (Table 1). However, some of the biases, such as undersampling or oversampling, the high-risk subpopulation affected (to some degree or another) some of the estimated epidemiological measures, in particular (understandably) the epidemic growth rate (Table 1). Importantly, none of the sampling biases affected appreciably the year of epidemic identification as estimated using the routine HIV testing, but these biases generally affected the year of epidemic identification as estimated using AIDS case notifications (Fig. 5).

4. Discussion

The aim of this study was to demonstrate, using mathematical modeling, the concept of using routine HIV testing data (such as those of blood donors (Hermez et al., 2010)) to identify a hidden HIV epidemic in some hard-to-reach high-risk subpopulation. We found that simple analyses on HIV testing data (of no methodological complexity) can provide a useful tool to identify such hidden epidemic, thereby supporting HIV surveillance efforts at a low cost that can be affordable even in resource-limited settings. We also found that this approach can identify the hidden epidemic rapidly, within at most few years of epidemic emergence, and much earlier than that of a surveillance based on AIDS case notifications. Use of routine HIV testing data can demonstrate to policymakers the emergence of an epidemic, thereby supporting advocacy for investment in HIV surveillance and intervention response.

Despite the limitations of HIV testing data, the presented analyses demonstrate that these data are of utility and can capture, despite some caveats and limitations, key temporal patterns of a hidden HIV epidemic in a hard-to-reach subpopulation. With availability of other supplementary data sources, such as HIV case notifications (which are routine across countries) (Bozicevic, Riedner, & Haghdoost, 2014; Kouyoumjian et al., 2013; Mumtaz et al., 2010, 2014b), and data on proxy biomarkers of risk behavior (such as prevalence of different sexually transmitted infections for sexual high-risk groups, or prevalence of hepatitis C virus for PWID) (Abu-Raddad et al., 2010c; Akbarzadeh et al., 2016; Mumtaz et al., 2015; Vickerman et al., 2010), the utility of this approach is enhanced and can be used to identify the specific high-risk subpopulation affected by the epidemic. It can also provide estimates for a wider set of epidemiological measures, such as high-risk subpopulation size estimates, growth rate of the epidemic in the high-risk subpopulation, and HIV epidemic potential.

An example to this end is the use of recently-notified HIV cases to identify the population affected by the epidemic, as suggested for the epidemic among MSM in Oman and Syria, and the epidemic among PWID in Saudi Arabia (Mumtaz et al., 2014b). Similarly, the presented approach can also supplement other sources of data, such as those of the integrated bio-behavioral surveillance surveys (Bozicevic et al., 2013), to provide a more comprehensive characterization of the epidemiology of HIV in a given population.

These findings are specially relevant in countries at low HIV prevalence, such as many of the countries in the MENA region (Abu-Raddad et al., 2010a, 2010b; Mumtaz et al., 2011, 2014a, 2014b), where there has been overall limited national commitment to HIV surveillance (Abu-Raddad, 2010; Abu-Raddad et al., 2010a; Bozicevic et al., 2013; Sawires et al., 2009), in

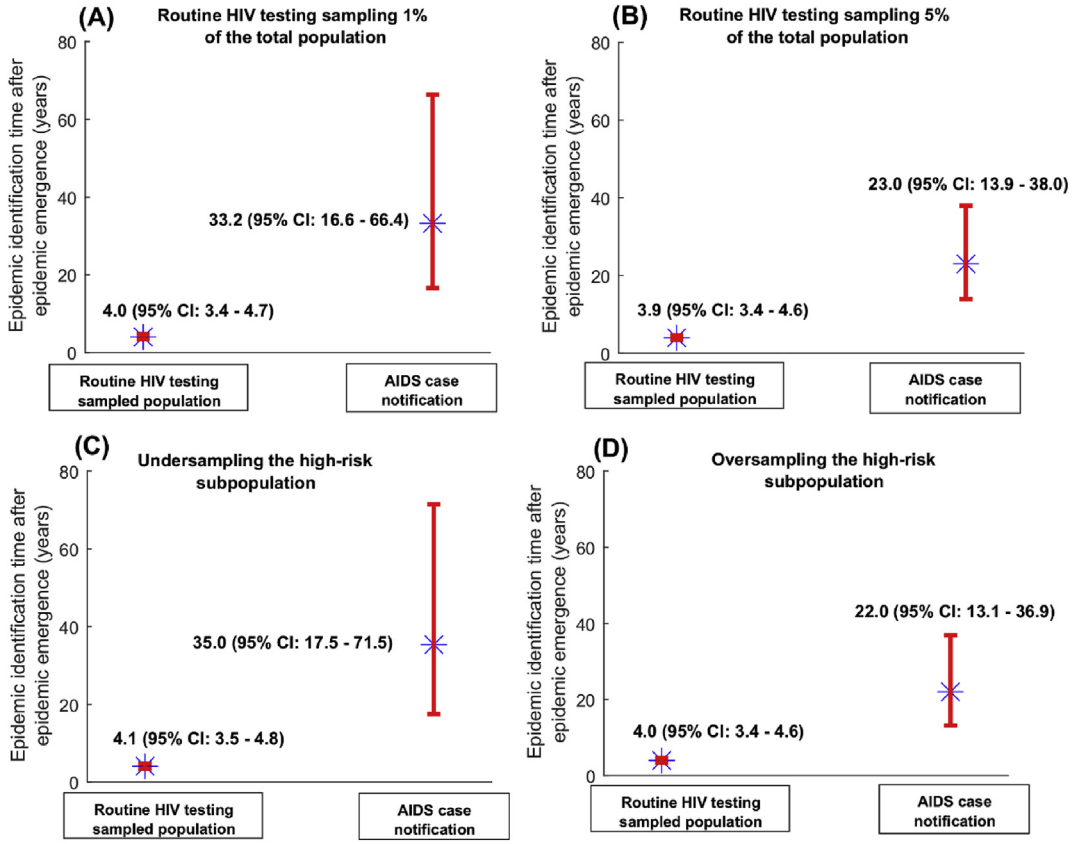

Fig. 5. Sensitivity analyses on the time of detecting the epidemic after epidemic emergence. A) Assuming a smaller sample size for the simulated routine HIV testing (1% of the total population). B) Assuming a larger sample size for the simulated routine HIV testing (5% of the total population). C) Assuming undersampling of the high-risk subpopulation in the simulated routine HIV testing. D) Assuming oversampling of the high-risk subpopulation in the simulated routine HIV testing.

part because of the high stigma associated with HIV (Abu-Raddad et al., 2010a; Sawires et al., 2009). The HIV epidemics in this region emerged mostly during the last decade and have not yet reached their peak (Abu-Raddad et al., 2010a, 2010b; Mumtaz et al., 2011, 2014a, 2014b). Existing evidence suggests considerable HIV epidemic potential in high-risk subpopulations in this region (Abu-Raddad et al., 2010c; Akbarzadeh et al., 2016; Mumtaz et al., 2015). Importantly, many of the epidemics, such as those among PWID in Libya and Pakistan (Mirzoyan et al., 2013; Mumtaz et al., 2014a), were discovered years after epidemic emergence (Mumtaz et al., 2011, 2014a, 2014b). Indeed, there is already anecdotal evidence that suggests ongoing emerging HIV epidemics in several countries which are yet to be documented, characterized, and tackled (Mumtaz et al., 2014b). With the approach proposed here, such potential epidemics could be detected early, overcoming some of the complexities and difficulties of directly monitoring epidemics in hidden and hard-to-reach subpopulations (Abu-Raddad, 2010; Abu-Raddad et al., 2010a, 2013).

Our study has limitations. We presented an approach for detecting emerging HIV epidemics, but other approaches could also be possible and should be investigated, such as using antiretroviral therapy (ART) programs to monitor trends—an approach that may not yet be feasible given the low ART coverage in MENA, the lowest globally at only 17% in 2015 (UNAIDS, 2016). While our proposed approach can provide qualitative characterization of hidden epidemics, it may only provide rough quantitative estimates for several of the epidemiologic measures. The provided estimates may not also be sufficient to establish a meaningful characterization of an emerging epidemic without other supplementary data, such as case notifications, to identify the affected population. Routine HIV testing data can be subject to multiple biases, and the nature of bias may vary by time and geographic location, as well as by country, thereby introducing challenges in generating epidemiological inferences.

The model used here was strictly designed to describe HIV sexual transmission, but the essential concept of this study relies merely on the generic existence of an emerging epidemic. Accordingly, the findings are applicable for epidemics among PWID. Indeed, the often rapidly-rising epidemics among PWID (Mumtaz et al., 2014a), are even easier to discern with this approach, as the growth rate of HIV prevalence is faster.

Epidemic dynamics can be complex whereby several epidemics could be emerging simultaneously in different high-risk subpopulations, further complicating the generation of epidemiological inferences. We did not factor ART coverage in the model, but this may not have affected our results, as we focused on emerging epidemics where ART coverage is yet to be scaled up.

Despite these limitations, our results suggest that this approach can be effective and easy to implement. The sensitivity analyses we conducted, to explore the impact of several such limitations, also supported the value of this approach.

5. Conclusions

Using hypothesis-generation modeling, we demonstrated the concept of using routine HIV testing data to identify and characterize hidden and emerging HIV epidemics in high-risk subpopulations. This approach can be easily implemented and requires limited resources, and can be applied alone, or along with other complementary data sources to provide a meaningful characterization of hidden epidemics. The approach also avoids challenges and difficulties of HIV surveillance among stigmatized, hidden, and hard-to-reach subpopulations. It would be pertinent as a next step to explore an application of this approach in a country that has sufficient data on high-risk subpopulations, along with different sources of routine HIV testing data, to empirically validate this concept and explore strengths and limitations of this approach. Lastly, our study demonstrates mathematical modeling as a powerful tool to investigate theoretical concepts for the development of programs and policies towards addressing difficult public health challenges and priorities.

Conflicts of interest

The authors have no competing interests to declare.

Financial support

This publication was made possible by NPRP grant number 9-040-3-008 from the Qatar National Research Fund (a member of Qatar Foundation). The findings achieved herein are solely the responsibility of the authors. The authors are also grateful for infrastructure support provided by the Biostatistics, Epidemiology, and Biomathematics Research Core at Weill Cornell Medicine-Qatar.

Authors' contributions

LJA conceived and led the design of the study and analyses. HHA and SFA contributed to model design, conducted simulations and analyses, and contributed to results interpretation. All authors contributed to drafting of the article.

Funding

The Qatar National Research Fund (NPRP 9-040-3-008).

Disclose funding received for this work

Others.

List of abbreviations

ART	Antiretroviral therapy
CI	Confidence interval
MSM	Men who have sex with men
PWID	People who inject drugs
FSWs	Female sex workers
MENA	Middle East and North Africa
SM	Supplementary Material
UNAIDS	Joint United Nations Programme on HIV/AIDS

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.idm.2018.10.001.

References

Abdel Messih, I. Y., Ismail, M. A., Saad, A. A., et al. (2014). The degree of safety of family replacement donors versus voluntary non-remunerated donors in an Egyptian population: A comparative study. *Blood Transfusion, 12*, 159–165.

Abu-Raddad, L. J. (2010 December 1). HIV prevention policy in the Middle East and North Africa: Entangled dilemmas. *Nature Middle East*.

Abu-Raddad LJ, Akala FA, Semini I, et al. Characterizing the HIV/AIDS epidemic in the Middle East and North Africa: Time for strategic action. Middle East and North Africa HIV/AIDS epidemiology synthesis project. World Bank/UNAIDS/WHO Publication. Washington DC: The World Bank Press; 2010.

Abu-Raddad, L. J., Hilmi, N., Mumtaz, G., et al. (2010). Epidemiology of HIV infection in the Middle East and North Africa. *AIDS, 24*(Suppl 2), S5–S23.

Abu-Raddad, L. J., & Longini, I. M., Jr. (2008). No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa. *AIDS, 22*, 1055–1061.

Abu-Raddad, L. J., Patnaik, P., & Kublin, J. G. (2006). Dual infection with HIV and malaria fuels the spread of both diseases in sub-Saharan Africa. *Science, 314*, 1603–1606.

Abu-Raddad, L. J., Schiffer, J. T., Ashley, R., et al. (2010). HSV-2 serology can be predictive of HIV epidemic potential and hidden sexual risk behavior in the Middle East and North Africa. *Epidemics, 2*, 173–182.

Abu-Raddad LJ, Sgaier SK, Mumtaz GR. HIV response in the Middle East and North Africa: An epidemic and its policy dilemmas. In: Smith R, editor. Global HIV/AIDS politics, policy and activism: Persistent challenges and emerging issues. Washington DC: Praeger Publishers; 2013. p. 143-168.

Akbarzadeh, V., Mumtaz, G. R., Awad, S. F., et al. (2016). HCV prevalence can predict HIV epidemic potential among people who inject drugs: Mathematical modeling analysis. *BMC Public Health, 16*, 1216.

Anderson R, May R. Infectious diseases of humans: Dynamics and control. Oxford University Press; 1991.

Awad, S. F., & Abu-Raddad, L. J. (2014). Could there have been substantial declines in sexual risk behavior across sub-Saharan Africa in the mid-1990s? *Epidemics, 8*, 9–17.

Awad, S. F., Sgaier, S. K., Tambatamba, B. C., et al. (2015). Investigating voluntary medical male circumcision program efficiency gains through subpopulation prioritization: Insights from application to Zambia. *PloS One, 10*. e0145729.

Berchenko, Y., & Frost, S. D. W. (2011). Capture-recapture methods and respondent-driven sampling: Their potential and limitations. *Sexually Transmitted Infections, 87*, 267–268.

Bozicevic, I., Riedner, G., & Calleja, J. M. (2013). HIV surveillance in MENA: Recent developments and results. *Sexually Transmitted Infections, 89*(Suppl 3), iii11–i16.

Bozicevic, I., Riedner, G., & Haghdoost, A. (2014). HIV case reporting in the countries of North Africa and the Middle East. *Journal of the International AIDS Society, 17*, 18962.

Caceres, C. F., Konda, K., Segura, E. R., et al. (2008). Epidemiology of male same-sex behaviour and associated sexual health indicators in low- and middle-income countries: 2003-2007 estimates. *Sexually Transmitted Infections, 84*(Suppl 1), i49–i56.

Diaz, T., Garcia-Calleja, J. M., Ghys, P. D., et al. (2009). Advances and future directions in HIV surveillance in low- and middle-income countries. *Current Opinion in HIV and AIDS, 4*, 253–259. https://doi.org/10.1097/COH.0b013e32832c1898.

Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems, 44*, 174–199.

Hermez, J., Petrak, J., Karkouri, M., et al. (2010). A review of HIV testing and counseling policies and practices in the Eastern Mediterranean region. *AIDS, 24*(Suppl 2), S25–S32.

Joint United Nations Programme on HIV/AIDS. Guidelines for second generation HIV surveillance: UNAIDS; 2000.

Kouyoumjian, S. P., El Rhilani, H., Latifi, A., et al. (2018). Mapping of new HIV infections in Morocco and impact of select interventions. *International Journal of Infectious Diseases, 68*, 4–12.

Kouyoumjian, S. P., Mumtaz, G. R., Hilmi, N., et al. (2013). The epidemiology of HIV infection in Morocco: Systematic review and data synthesis. *International Journal of STD & AIDS, 24*, 507–516.

Magnani, R., Sabin, K., Saidel, T., et al. (2005). Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS, 19*(Suppl 2), S67–S72.

May, R. M., Anderson, R. M., & Irwin, M. E. (1988). The transmission dynamics of human immunodeficiency virus (HIV) [and discussion]. *Philosophical Transactions of the Royal Society of London B Biological Sciences, 321*, 565–607.

McFarland, W., & Caceres, C. F. (2001). HIV surveillance among men who have sex with men. *AIDS, 15*(Suppl 3), S23–S32.

Mercer, C. H., Hart, G. J., Johnson, A. M., et al. (2009). Behaviourally bisexual men as a bridge population for HIV and sexually transmitted infections? Evidence from a national probability survey. *International Journal of STD & AIDS, 20*, 87–94.

Mills, S., Saidel, T., Magnani, R., et al. (2004). Surveillance and modelling of HIV, STI, and risk behaviours in concentrated HIV epidemics. *Sexually Transmitted Infections, 80*, ii57–ii62.

Ministry of Development Planning and Statistics. reportQatar's fourth national human development report: Realising Qatar national vision 2030, the right to development (Available at: http://www.gsdp.gov.qa/portal/page/portal/gsdp_en/knowledge_center/Tab2/NHDR4%20Complete%20Report%20English%20LowResolution%2028May2015.pdf): Ministry of Development Planning and Statistics 2015.

Ministry of Public Health. Personal communication. 2017.

Mirzoyan, L., Berendes, S., Jeffery, C., et al. (2013). New evidence on the HIV epidemic in Libya: Why countries must implement prevention programs among people who inject drugs. *JAIDS Journal of Acquired Immune Deficiency Syndromes, 62*, 577–583. https://doi.org/10.1097/QAI.0b013e318284714a.

Mumtaz, G. R., Awad, S. F., Feizzadeh, A., et al. (2018). HIV incidence among people who inject drugs in the Middle East and North Africa: Mathematical modelling analysis. *Journal of the International AIDS Society, 21*, e25102.

Mumtaz, G. R., Hilmi, N., McFarland, W., et al. (2011). Are HIV epidemics among men who have sex with men emerging in the Middle East and North Africa?: A systematic review and data synthesis. *PLoS Medicine, 8*. e1000444.

Mumtaz G, Hilmi N, Zidouh A, et al. HIV mode of transmission analysis in Morocco 2010: Morocco Ministry of health & Joint United Nations Programme on HIV/AIDS; 2010.

Mumtaz, G. R., Kouyoumjian, S. P., Hilmi, N., et al. (2013). The distribution of new HIV infections by mode of exposure in Morocco. *Sexually Transmitted Infections, 89*. iii49-iii56.

Mumtaz, G. R., Riedner, G., & Abu-Raddad, L. J. (2014 Mar). The emerging face of the HIV epidemic in the Middle East and North Africa. *Current Opinion in HIV and AIDS*, 183–191.

Mumtaz, G. R., Weiss, H. A., Thomas, S. L., et al. (2014). HIV among people who inject drugs in the Middle East and North Africa: Systematic review and data synthesis. *PLoS Medicine, 11*. e1001663.

Mumtaz, G. R., Weiss, H. A., Vickerman, P., et al. (2015). Using hepatitis C prevalence to estimate HIV epidemic potential among people who inject drugs in the Middle East and North Africa. *AIDS, 29*, 1701–1710.

Pinkerton, S. D. (2008). Probability of HIV transmission during acute infection in Rakai, Uganda. *AIDS and Behavior, 12*, 677–684.

Rehle, T., Lazzari, S., Dallabetta, G., et al. (2004). Second-generation HIV surveillance: Better data for decision-making. *Bulletin of the World Health Organization, 82*, 121–127.

Sawires, S., Birnbaum, N., Abu-Raddad, L. J., et al. (2009 Jul 1). Twenty-five years of HIV: Lessons for low prevalence scenarios. *Journal of Acquired Immune Deficiency Syndromes*, S75–S82.

Suligoi, B., Raimondo, M., Regine, V., et al. (2010). Epidemiology of human immunodeficiency virus infection in blood donations in Europe and Italy. *Blood Transfusion, 8*, 178–185.

UNAIDS. Epidemiological software and tools. http://www.unaids.org/en/KnowledgeCentre/HIVData/Epidemiology/epi_software2007.asp. 2008.

UNAIDS. Global AIDS update 2016. Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland. Found at http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf. Last accessed October 2016. 2016.

UNAIDS/WHO. AIDS epidemic update 2010: UNAIDS fact sheet (available at: http://www.unaids.org/documents/20101123_FS_SSA_em_en.pdf, accessed 23 July 2012)2010a.
UNAIDS/WHO. Epidemiological data, HIV estimates 1990-2009 (available at:: http://www.unaids.org/en/dataanalysis/epidemiology/)2010b.
Vickerman, P., Hickman, M., May, M., et al. (2010). Can hepatitis C virus prevalence be used as a measure of injection-related human immunodeficiency virus risk in populations of injecting drug users? An ecological analysis. *Addiction, 105*, 311–318.
Walker, N., Stover, J., Stanecki, K., et al. (2004). The workbook approach to making estimates and projecting future scenarios of HIV/AIDS in countries with low level and concentrated epidemics. *Sexually Transmitted Infections, 80*, i10–i13.
Wawer, M. J., Gray, R. H., Sewankambo, N. K., et al. (2005). Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *The Journal of Infectious Diseases, 191*, 1403–1409.
World Health Organization. Blood safety and availability. 2012 [updated June 2012; cited 2013 May 01]; Available from: http://www.who.int/mediacentre/factsheets/fs279/en/.