

Design and Development of an Automated Surveillance System for Outbreak Detection and Individual Risk Assessment for HIV and Viral Hepatitis B and C Among People Who Use Drugs: The Hippocrates Project

Touloumi G¹, Baralou V¹, Pantazis N¹, Karakosta A¹, Anagnostou O², Katsiris D³, Micha K², Pistikos G³, Danopoulos C², Rountsos K³, Theocharis A²

¹ Department of Hygiene, Epidemiology & Medical Statistics, Medical School, National and Kapodistrian University of Athens, Athens, Greece; ² Organization Against Drugs (OKANA); ³ InDigital

Background

- People who use drugs (PWUD) are at high risk of acquiring HIV, HBV and HCV.
- During approximately the last decade, several HIV outbreaks emerged among PWUD in many European countries and the USA, often preceded by undetected HCV outbreaks. ➡ Need for timely outbreak detection and identification of high-risk individuals.

Objective: To develop an automated system for real-time outbreak detection and an algorithm for individual risk assessment for HIV, HBV and HCV among PWUD. For that, an integrated information system for PWUD registration and monitoring was also developed.

Methods

Real-time detection of the start and subsequent epidemic states of an outbreak

Data

- Various data sources provided by the Greek Organization against Drugs (OKANA) within the “Hippocrates” project were combined to monitor new weekly diagnoses (Y_t) of each infection and region (Figure 1).

Statistical methods

- To detect the transition from the pre-epidemic state ($Z_t = 1$) to growth ($Z_t = 2$), we applied in real time:

- Control charts: Cumulative Sum (CUSUM), Exponentially weighted moving average (EWMA) chart

- A new method based on prediction interval (PI) (Figure 2)

- A novel approach, called HMM-PI, that combines a two-state hidden Markov model (HMM) with the PI method

$$\begin{pmatrix} w_t(1 - p_{12}) & 1 - w_t(1 - p_{12}) \\ 0 & 1 \end{pmatrix}, w_t = \begin{cases} \exp\left\{10\left(\frac{U}{Y_t} - 1\right)\right\}, & Y_t > U \\ 1, & Y_t \leq U \end{cases}$$

and

$$Y_t \sim NB(\mu_t, r_t), \text{ where } \mu_{t|Z_t=1} = \mu_{t-1} \text{ and } \mu_{t|Z_t=2} = \mu_{t-1}e^{\beta_1}, \beta_1 > 0$$
$$r_{t|Z_t=1} = r_1 \text{ and } r_{t|Z_t=2} = r_2 \text{ such that } r_1 < r_2$$

- To detect the subsequent epidemic states (i.e., plateau and/ or decline, post-epidemic state), PI and HMM-PI were properly modified.

Performance evaluation

- Sensitivity: true positive rate
- Specificity: true negative rate
- Timeliness: difference between the start of the epidemic and the first alarm signal after its onset

Data

- Data sources
 - Individual data: questionnaire TDI-DRID, laboratory test results
 - Aggregated data: harm reduction programs
- Inclusion criteria: year of diagnosis ≥ 2009 and infection date at least 30 days after the date of entrance in the program
- Baseline: date of entrance in the program
- Outcome: HIV, HBV and/ or HCV infection within one year from baseline
- Risk factors: demographic, socio-economic, behavioral, laboratory data
- Data were randomly split into training and test set with a ratio of 80:20.

Statistical methods

- Handling of missing data: multiple imputation by chained equations
- Variable selection: variable importance by Random Survival Forest (RSF)
- Final prognostic model: Cox regression using the risk factors selected
- Performance metric: time-dependent area under curve (AUC) on test set

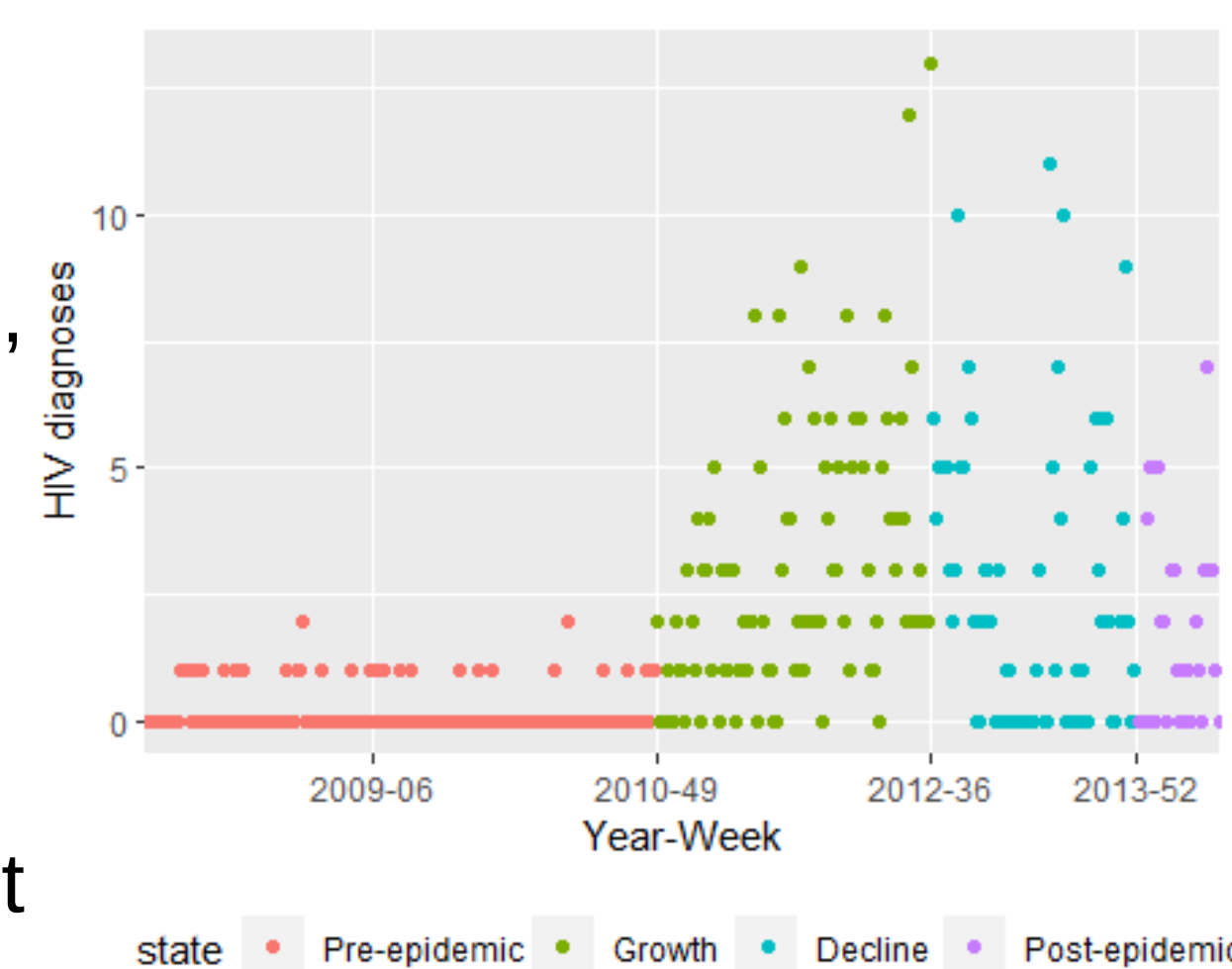


Figure 1. Weekly HIV cases among PWID in Athens segmented in four epidemic states

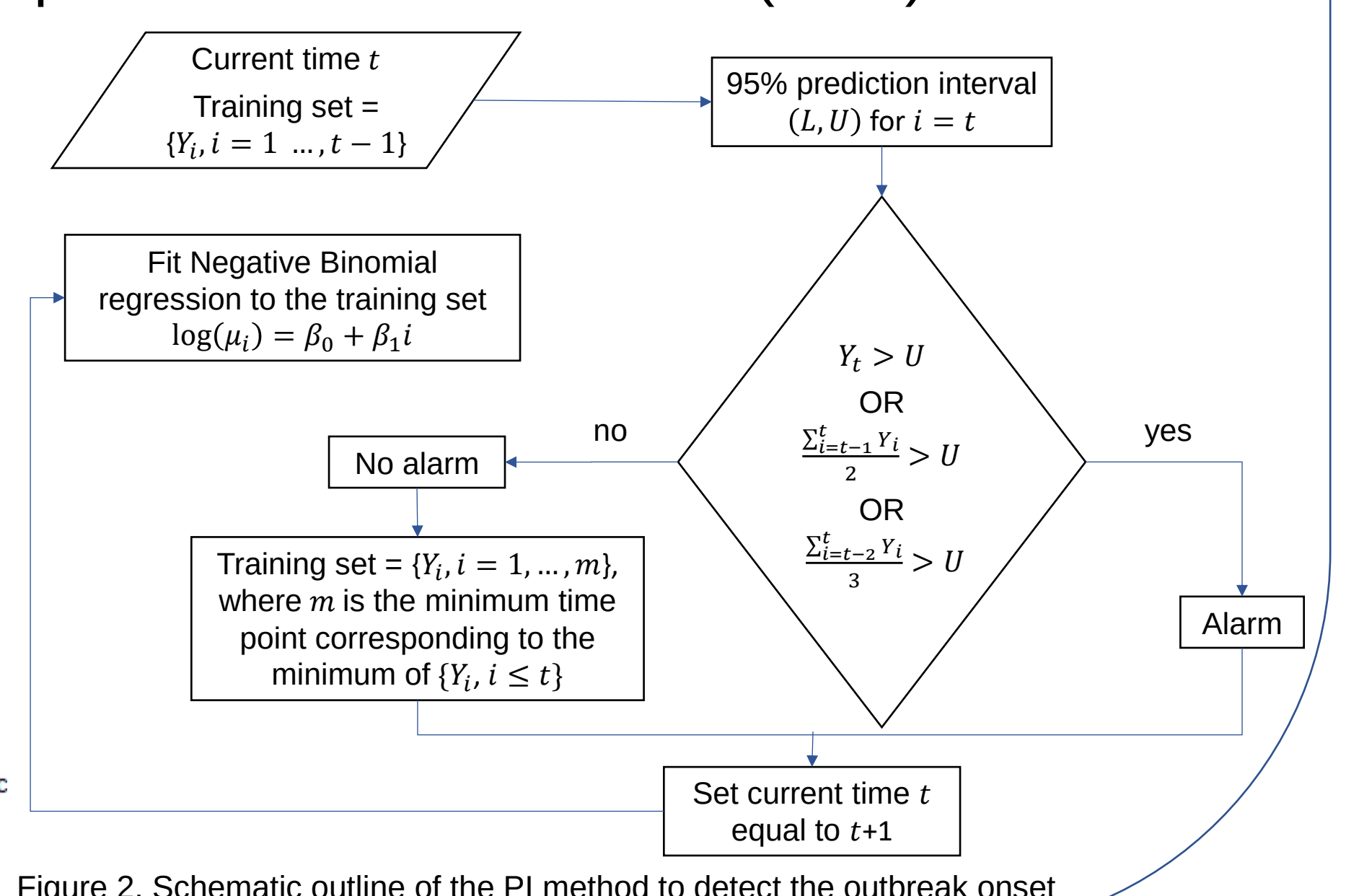


Figure 2. Schematic outline of the PI method to detect the outbreak onset

Results

System for real-time detection of each epidemic state

- The application is briefly presented in Figure 3. Five parameters are given as input by the user. Indicatively, the output is shown in Figure 4.
- Growth: excellent performance by all methods (sensitivity > 95%, specificity > 85%, no delay in alarm signal).
- Decline: HMM-PI had the highest performance (sensitivity 84%, specificity 89%, timeliness 15 weeks).
- Post-epidemic state: unsatisfactory performance.

Tool for individual risk assessment

- Nine factors were included as questions (Table).
- The final prognostic model gave satisfactory AUC (>70%) on test data.

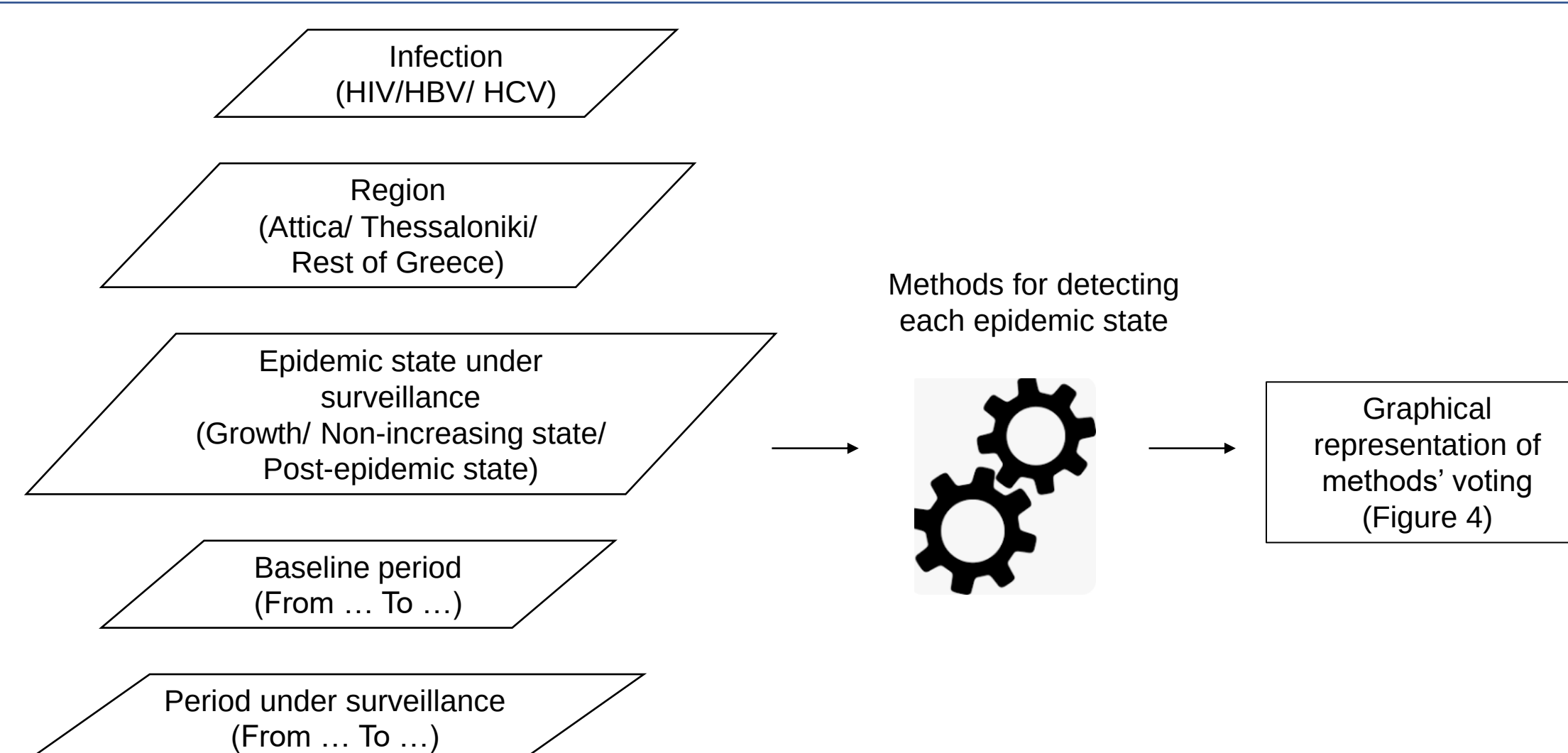


Figure 3. Schematic outline of the application created to monitor new diagnoses

Variable	HR (95% CI)
Age (years)	1.03 (1.02 – 1.03)
Permanent (stable) housing (Yes/No)	0.80 (0.51 – 1.03)
Unemployment (Yes/No)	1.22 (1.05 – 1.43)
Education level	
>6 years and ≤12 years vs. ≤6 years	0.95 (0.81 – 1.10)
>12 years vs. ≤6 years	0.71 (0.53 – 0.95)
Primary method of use being injection (Yes/No)	1.25 (1.04 – 1.50)
Daily use in the past month (Yes/No)	1.22 (1.05 – 1.43)
Injecting use in the past month (Yes/No)	1.30 (1.15 – 1.42)
Age at first injection (years)	0.97 (0.96 – 0.98)
Syringe sharing in the past month (Yes/No)	1.36 (1.26 – 1.45)

Table. Hazard ratios (HR) and 95% Confidence Intervals (CI) estimated by the Cox model for the risk factors selected by RSF

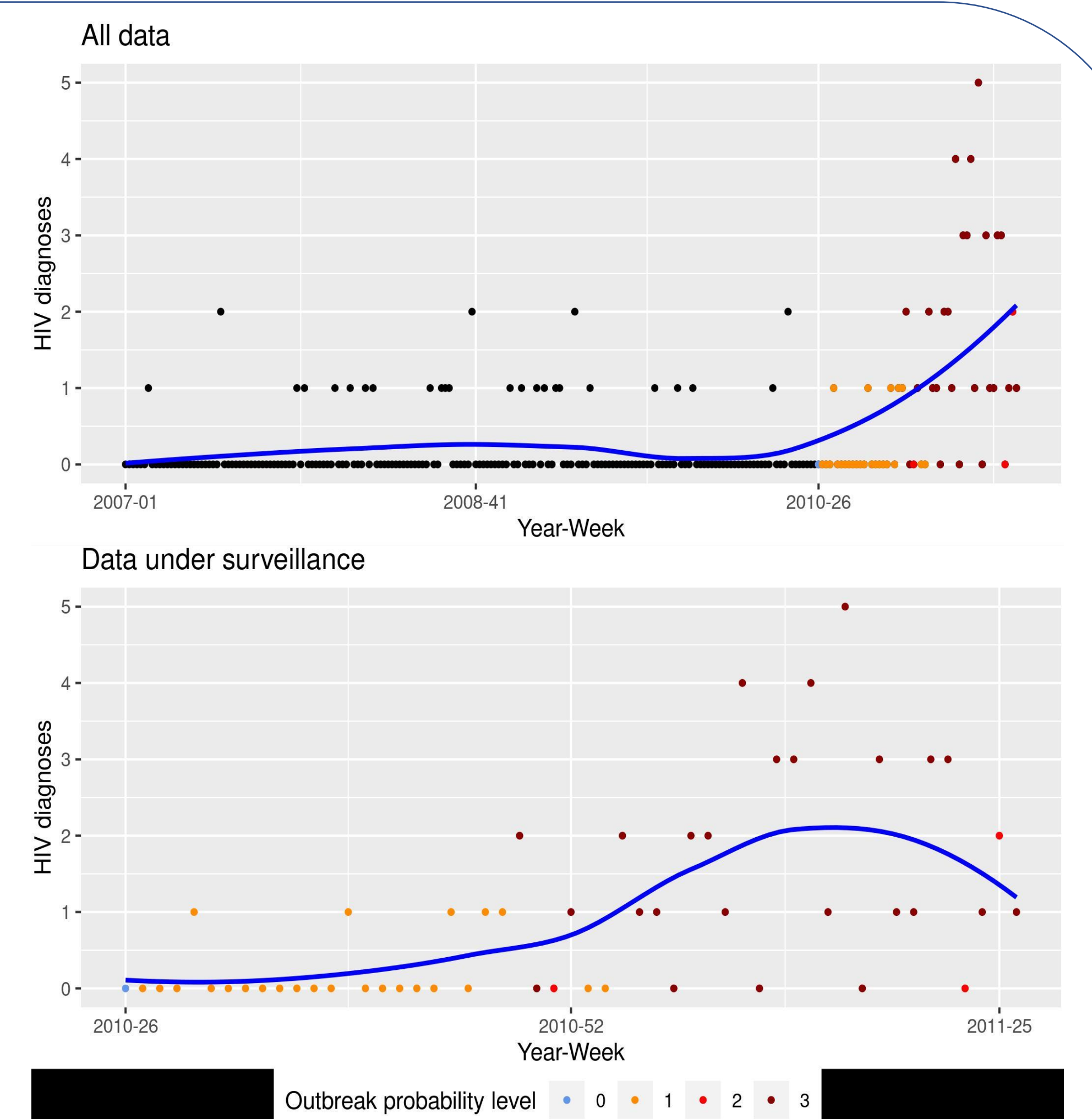


Figure 4. Output of the application to monitor HIV diagnoses in Attica. Epidemic state under surveillance: Growth. Baseline period from 2007-01 until 2010-25. Period under surveillance from 2010-26 until 2011-25.

Conclusions

- An automated system that monitors new HIV, HBV and HCV diagnoses among PWUD is an effective investment which can contribute to the timely control of an outbreak and the assessment of control measures.
- An automated tool that provides alert for high-risk individuals can help prevent new infections and enhance their awareness about the risk they face.
- Implementing the surveillance system “Hippocrates” in clinical practice can benefit not only PWUD but also other vulnerable populations susceptible to these infections.