Development of automated HIV case reporting system using national electronic medical record in Thailand

Thitipong Yingyong,[1] Suchunya Aungkulanon [ID] ,[2] Wasun Saithong,[3] Supiya Jantaramanee,[1] Kanjanakorn Phokhasawad,[2] Ian Fellows,[4] Thananda Naiwatanakul,[2] Jariya Mobnarin,[3] Narong Charoen,[3] Paiboon Waikayee,[3] Sanny Chen Northbrook[2]

¹Division of Epidemiology, Department of Disease Control, Ministry of Public Health, Nonthaburi, Thailand
²Division of Global HIV and TB, US Centers for Disease Control and Prevention, Nonthaburi, Thailand
³Information Communication Technology Center, Permanent Secretary Office, Ministry of Public Health, Nonthaburi, Thailand
⁴Division of Global HIV and TB, US Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Correspondence to
Dr Suchunya Aungkulanon; ooz2@cdc.gov

ABSTRACT

Background An electronic medical record (EMR) has the potential to improve completeness and reporting of notifiable diseases. We developed and assessed the validity of an HIV case detection algorithm and deployed the final algorithm in a national automated HIV case reporting system in Thailand.

Methods The HIV case detection algorithms leveraged a combination of standard laboratory codes, prescriptions and International Classification of Diseases, 10th Revision diagnostic codes to identify potential cases. The initial algorithm was applied to the national EMR from 2014 to June 2020 to identify HIV-infected subjects to build the national HIV case reporting system (Epidemiological Intelligence Information System (EIIS)). A subset of potential positives identified by the initial algorithm were then validated and reviewed by infectious disease specialists. This review identified that a proportion of the false positives were due to pre-exposure prophylaxis/postexposure prophylaxis (PrEP/PEP) antiretrovirals, and so the algorithm was refined into a 'Final Algorithm' to address this.

Results Positive predictive value of identifying HIV cases was 90% overall for the initial algorithm. Individuals misclassified as HIV-positive were HIV-negative patients with incorrect diagnostic codes, prescription records for PrEP, PEP and hepatitis B treatment. Additional revision to the algorithm included triple drug regimen to avoid further misclassification. The final HIV case detection algorithm was applied to national EMR between 2014 and 2020 with 449 088 HIV-infected subjects identified from 1496 hospitals. EIIS was designed by applying the final algorithm to automated extract HIV cases from the national EMR, analysing them and then transmitting the results to the Ministry of Public Health.

Conclusions EMR data can complement traditional provider-based and laboratory-based disease reports. An automated algorithm incorporating laboratory, diagnosis codes and prescriptions have the potential to improve completeness and timeliness of HIV reporting, leading to the implementation of a national HIV case reporting system.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Traditional passive surveillance is burdensome to clinicians, and it is often incomplete and delayed as it may lack information needed for public health purposes.
⇒ Electronic medical record (EMR) is capable of accurately reporting and monitoring notifiable diseases.

WHAT THIS STUDY ADDS

⇒ This study developed HIV case detection algorithms using the national EMR database which validated using National AIDS Programme data and reviewed by infectious disease specialists.
⇒ Our novel algorithm to identify HIV infection in the EMR system contributed to the development of an automated HIV case reporting system in Thailand.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ HIV case detection algorithms have the potential to improve completeness and timeliness of HIV reporting system compared to the traditional system.
⇒ Much effort should be concentrated on improving data quality of national EMR data.

BACKGROUND

Electronic medical record (EMR) has the potential to improve completeness and reporting of notifiable diseases beyond traditional clinician-initiated and laboratory-based disease reporting systems.[1] Traditional passive surveillance is burdensome to clinicians, and it is often incomplete and delayed as it may lack information needed for public health purposes (eg, patient signs and symptoms, prescribed treatments and pregnancy status).[2 3] EMR, however, contains this information and stores it in a form that can be used for electronic analysis and reporting. Consequently, EMR-based reporting has the potential to provide active notifiable disease surveillance that is more timely, complete and clinically detailed that enables longitudinal disease reporting and analysis. With the advent and adoption of EMR, researchers are now able to rapidly identify potential disease

cases for clinical studies. Disease detection algorithms are needed to search across billing data, laboratory data and clinical documentation to perform case detection. These disease detection algorithms can be conceived in a manner that has high sensitivity and specificity for identifying individual's true disease status using methods borrowed from routine clinical care.[4]

The first case of HIV/AIDS in Thailand was officially reported in July 1984.[5] Shortly after, HIV/AIDS was declared a highly infectious disease requiring mandatory notification in 1985. In the past, physicians were required to report all patients with HIV infection, including asymptomatic cases in Report 506/1 to the provincial health office, who forwarded the information to the Ministry of Public Health (MoPH). This passive surveillance relies on physicians to report new cases of HIV infection or AIDS directly to the MoPH. Data from passive surveillance are often slow to accrue and incomplete and may not support a timely and well-aimed public health purpose. To address the problem MoPH stopped mandating asymptomatic HIV case reporting using Report 506/1 in 2014 and classified HIV as a notifiable condition and mandated healthcare providers and laboratories to report HIV cases to provincial health officials in the Communicable Disease Act BE 2558 (2015).[6]

HIV infection is a disease that lends itself to an algorithm-based case detection, given the reliance on laboratory-based testing. Previous studies have attempted to identify HIV-infected patients in selected population or selected hospitals in developed countries.[7–10] However, there is no validated procedure for using data from medical records to identify diagnosed HIV-infected patients for national HIV/AIDS surveillance purposes. In order to maximise use of EMR for automated HIV/AIDS reporting system, we developed and assessed the validity an HIV case detection algorithm and estimated the positive predictive value (PPV) of the algorithm to detect new diagnosed HIV cases and develop a national automated HIV case reporting system in Thailand.

EMR database

Since 2007, Thailand MoPH has established a national electronic health system to house the central and provincial health data centres (HDC) for management of EMR. All levels of health facilities under the MoPH are required to upload and transfer disaggregated, individualised data to a central database using cloud technology at least once a month. In 2021, the MoPH HDC platform received EMR data (ie, patient demographics, vital signs, test orders, test results, prescriptions, diagnostic codes and healthcare provider details) from 947 MoPH public hospitals, 55 non-MoPH public hospitals and 9760 subdistrict health promotion hospitals. Thai healthcare is dominated by public health facilities accounting for 79% of hospital beds[11] and with 21% of total beds in private hospitals. Specialised HIV treatment and care services are mainly provided in public health facilities under the

management of universal coverage (UC) to ensure equitable access.

The standard patient-level data collected at each health facility includes demographics and health services.[12] Data are subsequently managed and summarised for key performance indicators. The visualised report is accessible for health management and disease control on the web-based HDC dashboard (http://hdcservice.moph.go.th)

METHODS

Initial algorithm development

HIV infection was identified by applying the surveillance case definition of HIV/AIDS from the 2015 national guidelines for reporting notifiable communicable disease in Thailand.[13] In order to maximise the utility and performance characteristics of the algorithm, six separate conditions using complementary approaches to HIV case detection were created. A panel of physicians, including experts in the diagnosis and treatment of HIV infection, provided recommendations for the development of these conditions. The HIV case detection algorithm combined national laboratory testing codes and results, Thai Medicine Terminology prescriptions[14] and International Classification of Diseases, 10th Revision (ICD-10) diagnostic codes. A list of laboratory tests, ICD-10 and antiretroviral (ARV) medication can be found in online supplemental appendix 1.

We identified six conditions under which a classification of HIV positive would be a reasonable conclusion and ordered them based on the panel's likelihood of false positivity. The initial algorithm classified an individual as positive if any one of the six conditions (A, B1, B2, C1, C2 and D1) were met. Condition A is met if laboratory test (anti-HIV, HIVDNA PCR or viral load) identified the individual as HIV infected. For HIV antibody testing, only a 'positive' result was defined as positive, whereas 'negative' and 'inconclusive' results were defined as negative. The presence of detectable viral load was defined as HIV-positive regardless of test result. Conditions B1, B2, C1 and C2 were based on clinical evidence and designed to detect those patients not identified by condition A due to incomplete HIV testing results in EMR. As a result of confidentiality concerns with sharing HIV laboratory results, the laboratory information system not linked to the hospital information system (HIS), and in other cases formatting issues when transferring data from hospital HIS to MoPH HDC prevent HIV laboratory results from being represented in the EMR. Condition B1 was considered met if an individual had three or more HIV-related ICD10 events as well as ARV drug use for at least three visits. Condition B2 was considered satisfied if there were one or two HIV-related ICD10 events and a CD4 test result of <200. C1 is a weaker condition that is satisfied with the presence of any HIV-related ICD10 event in the patient history. C2 and D1 are satisfied based solely on ARV use with C1 triggered when ARVs are reported on

Table 1 Summary of criteria included in the initial HIV case detection algorithm

Surveillance criteria	Condition	Laboratory	Diagnosis	Pharmacy
Laboratory-confirmed evidence	A	Presence of positive HIV antibody or HIV DNA PCR (in children <12 months) or result of detectable viral load		
Presumptive clinical evidence	B1		Presence of at least three HIV-related ICD-10 events	Report of antiretroviral drugs for at least three visits
	B2	Result of CD4 test result <200 cell count	Presence of one or two HIV-related ICD-10 event	
Probable clinical evidence	C1		Presence of at least one HIV-related ICD-10 event	
	C2			Report of antiretroviral drugs for at least two clinic visits, dated more than 60 days apart
Possible clinical evidence	D1			Report of antiretroviral drugs for two clinic visits, dated less than or equal to 60 days apart

ICD-10, International Classification of Diseases, 10th Revision.

two visits >60 days apart and D1 is triggered if they are ≤60 days apart. The initial algorithm distinguishes four levels of confidence in a determination of a positive, ranging from laboratory confirmed to possible clinical evidence. Our algorithm criteria are summarised in table 1.

Assessing validity
We assessed the validity the initial algorithm by applying it to the MoPH EMR from 31 hospitals that were interested in participating from 3 high HIV burden provinces in 2017. HIV positive patients from 31 hospitals were extracted from 460,575 HIV positives cases in the MoPH HDC platform to assess algorithm performance. To verify that subjects were HIV-infected, we used national AIDS programme (NAP) data for 'true HIV-positive' and cross-matched with individuals identified through the HIV case detection algorithm. The NAP is National Health Security Office's electronic database used for recording clinical and laboratory services for HIV monitoring and reimbursement.[15] Identification of an HIV-infected patient by the NAP occurs at the hospital, where local health-care providers register all HIV patients into the NAP for HIV test, CD4, and viral load test reimbursement for all Thai citizens regardless of health insurance scheme. This helps ensure that patients identified by the NAP are truly infected with HIV. Charts of patients who were not identified by NAP were further reviewed by infectious disease specialists from that hospital. The PPV of identifying HIV/AIDS cases was used to measure accuracy of the algorithm. We also reviewed the distribution of subjects by number of diagnostic codes and confirmatory

evidence of HIV infection using HIV viral load and ARV therapy results.

Final algorithm development
The expert reviews for the false positives found in the validation step were then analysed to determine whether any refinement to the algorithm conditions could be made. Frequencies were calculated for the expert determined reason for the misclassification and each reason was assessed to investigate whether there were additional constraints to the conditions that could be added to reduce the false positive likelihood. The resulting algorithm was labelled the final algorithm and was applied to the MoPH EMR from 2014 to 2020 to identify HIV-infected subjects for HIV case reporting system.

Automated HIV case reporting system development
The final algorithm was used to build the new HIV case reporting system called Epidemiological Intelligence Information System (EIIS). Data were integrated, stored, managed and analysed in the MoPH HDC cloud-based warehouse. EIIS Data warehouse security measures were strictly enforced to ensure data integrity at all levels. The system limited access to different level of authorised users. The data warehouse was set up to be read-only by default to prevent any threatening from being executed on the data. National identification number was encrypted applying the customised Hash Algorithm. Specifically, only the hashed version of national ID was stored in a database, which was decrypted to only authorised users.

RESULTS
Application of the initial HIV case detection algorithm

We applied the initial algorithm to the MoPH EMR data. All patients who were seen at hospitals under the MoPH HDC platform between 2014 and 2020 were included in this study. The HIV case detection algorithm identified a total of 460 575 cases among all patients receiving HIV services in 947 MoPH hospitals, 2243 subdistrict health promotion hospitals and 55 non-MoPH public hospitals. Twelve per cent of the cases were reported to have died. Table 2 shows unduplicated counts of patient who met the case detection criteria for each condition. Individuals

Table 2 Demographic characteristics of identified HIV cases by condition, 2014–2020

	A		B1		B2		C1		C2		D1	
	N	%	N	%	N	%	N	%	N	%	N	%
Total individuals	11 487		337 453		838		91 256		6543		12 998	
Sex												
Male	6572	57.2	186 215	55.2	574	68.5	54 780	60	3270	50	3658	28.1
Female	4915	42.8	151 238	44.8	264	31.5	36 475	40	3273	50	9340	71.9
Health insurance												
UC	7881	68.6	237 534	70.4	509	60.7	51 762	56.7	3048	46.6	6046	46.5
CSMBS	447	3.9	11 228	3.3	29	3.5	5133	5.6	1200	18.3	1922	14.8
SSS	2010	17.5	58 418	17.3	200	23.9	14 157	15.5	1445	22.1	2344	18
Migrant	271	2.4	10 975	3.3	22	2.6	6451	7.1	167	2.6	843	6.5
Self-paid	878	7.6	19 298	5.6	78	9.3	13 753	14.8	683	10.2	1843	13.9
Age												
0–15 years	136	1.2	3360	1	5	0.6	6044	6.6	231	3.5	4269	32.8
16–20 years	360	3.1	7702	2.3	22	2.6	4102	4.5	134	2.1	1221	9.4
21–25 years	765	6.7	18 340	5.4	99	11.8	8877	9.7	873	13.3	2617	20.1
26–30 years	889	7.7	23 776	7.1	101	12.1	9901	10.9	888	13.6	1663	12.8
31–35 years	1047	9.1	29 300	8.7	120	14.3	10 675	11.7	526	8	1000	7.7
36–40 years	1453	12.7	42 831	12.7	125	14.9	11 575	12.7	520	8	750	5.8
41–45 years	2013	17.5	61 388	18.2	131	15.6	10 995	12.1	556	8.5	537	4.1
46–50 years	1662	14.5	52 204	15.5	82	9.8	7598	8.3	483	7.4	283	2.2
>50 years	3162	27.5	98 552	29.2	153	18.3	21 489	23.6	2332	35.6	658	5.1
Nationality												
Thai	11 151	97.1	327 273	97	803	95.8	83 559	91.6	6450	98.6	12 687	97.6
Cambodia	14	0.1	916	0.3	1	0.1	879	1	12	0.2	37	0.3
Laos	21	0.2	1132	0.3	2	0.2	896	1	10	0.2	20	0.2
Myanmar	168	1.5	4858	1.4	15	1.8	3964	4.3	35	0.5	106	0.8
Others	133	1.2	3274	1	17	2	1958	2.2	36	0.6	148	1.1
Year of registration in EIIS												
2014	0	0	129 267	38.3	0	0	11 887	13	987	15.1	837	6.4
2015	0	0	70 436	20.9	0	0	13 267	14.5	979	15	1228	9.5
2016	0	0	41 094	12.2	0	0	10 798	11.8	1163	17.8	1586	12.2
2017	177	1.5	30 704	9.1	8	1	13 140	14.4	1262	19.3	2085	16
2018	2766	24.1	30 923	9.2	157	18.7	14 070	15.4	1045	16	2487	19.1
2019	4785	41.7	21 148	6.3	330	39.4	13 818	15.1	749	11.5	2736	21.1
2020	3759	32.7	13 881	4.1	343	40.9	14 276	15.6	358	5.5	2039	15.7

A1: Laboratory-confirmed evidence: HIV-positive test result or VL result; B1: Presumptive clinical evidence: at least 3 HIV-related ICD-10 events and report of ARV; B2: Presumptive clinical evidence: CD4 <200 and one or two HIV-related ICD-10 event; C1: Probable clinical evidence: at least one HIV-related ICD-10 event; C2: Probable clinical evidence: ARV use dated more than 60 days apart; D1: Possible clinical evidence: ARV use dated less than 60 days apart.

ARV, antiretroviral; CSMBS, Civil Servants' Medical Benefit Scheme; EIIS, Epidemiological Intelligence Information System; ICD-10, International Classification of Diseases, 10th Revision; SSS, social security scheme; UC, universal coverage.

were assigned to the highest evidence level for which they qualified. Seventy percent of cases were identified by both diagnosis and prescription (condition B1). Four percent of identified cases had strong evidence of HIV laboratory results (condition A) because HIV laboratory codes were included in the standard 43 files later on in 2017.

Accuracy of the initial algorithm

We identified 26 138 individual patients who met the HIV-infected case detection algorithms at 31 participating hospitals. After matching individual data with NAP, we found 18 647 records (71%) registered in NAP and an additional 7491 patients not identified by NAP. Patient charts of 7491 (100%) were reviewed by infectious disease specialists from their respective hospitals. Of these 7491, 4924 (66%) were correctly classified as HIV positive by our case-detection algorithms, 2120 (28%) were misclassified as HIV-positive (false positive) and 447 (6%) were not found in the hospital database. Positive cases who were not registered in NAP were mainly non-Thai citizens, covered under non-UC health insurance schemes or self-paid patients. PPV of identifying HIV cases based on the algorithm was (18 647+4924)/26 138=90% overall and 98% (16741+3481)/20701 among individuals receiving services in the past year.

Generating a final algorithm

Of the 2120 false positives identified by the review, the experts noted incorrect ICD10 code as the reason in 71% of cases where a reason was given (see table 3). This indicates that condition C1 might reduce the false positivity rate considerably; however, this would not be desirable as it also accounts for a large proportion of the individuals the algorithm classifies as positive (see table 2). As such, we felt that the cost in terms of the introduction of false negatives would be too high.

Postexposure prophylaxis (PEP) and pre-exposure prophylaxis (PrEP) accounted for 28% of false positives where a reason was given and Hepatitis B for an additional 1%. Recognising that (1) some medications used to treat HIV infection are also used for hepatitis B; (2) monotherapy of ARV drugs is not considered an effective HIV medication unless prescribed in combination with an additional drug and (3) dual therapy with emtricitabine and tenofovir are prescribed for PrEP and PEP, we further refined the
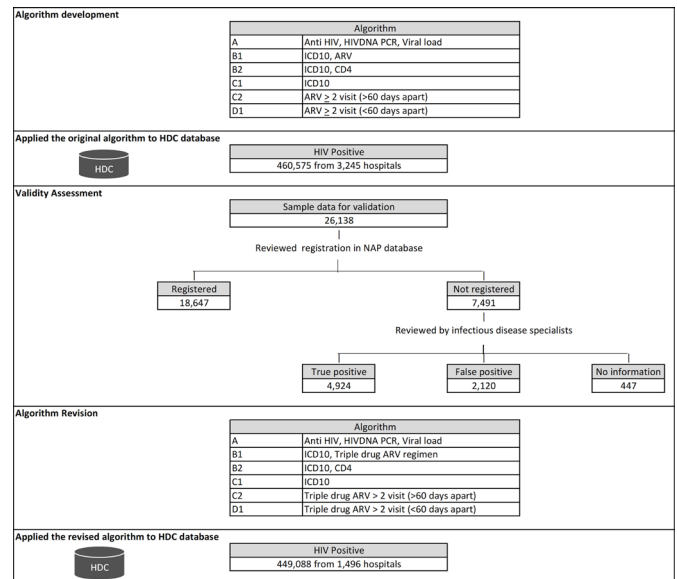
Table 3 Reasons given for false positivity by specialist review

	Count	%
PEP	415	26
PrEP	27	2
Hepatitis B	20	1
Incorrect ICD-10	1152	71
No information	506	

ICD-10, International Classification of Diseases, 10th Revision; PEP, postexposure prophylaxis; PrEP, pre-exposure prophylaxis.

Figure 1 Development and validity assessment of the HIV case detection algorithms. ARV, antiretroviral; HDC, health data centres; ICD-10, International Classification of Diseases, 10th Revision; NAP, National AIDS Programme.

algorithm to include triple drug regimen for HIV medication to avoid misclassification.

Figure 1 summarises development and validity assessment of the HIV case detection algorithms. The final algorithm replaces conditions B1, C2 and D1 by B1F, C2F and D1F, where the only difference is that the conditions for the final algorithm require triple drug ARV therapy to be triggered. Applying it to the MoPH EMR data found that the final algorithm classified 449 088 individuals as HIV positive, vs 460 575 for the initial algorithm. The similarity of volume of classified cases indicates that the restriction of the conditions to triple therapy was not overly restrictive.

The final algorithm experienced a large reduction in the number of hospitals, going from 3245 in the initial algorithm to 1496 in the final. Around 70% of hospitals are subdistrict health promoting hospitals providing primary care and health promotion services, but not ART. Triple drug regimen requirement reduced number of cases from these hospitals. Thus it is unsurprising that most cases in the EMR system should come from a more limited number of hospitals, whereas PrEP and PEP regimens would be present in a wider array of facilities.

DISCUSSION

To our knowledge, this is the first description of an algorithm using EMR data to identify patients with HIV positive status in Thailand. Our findings are consistent with the NAP. In June 2021, the final algorithm as implemented in the EIIS identified a total of 517 503 cases compared with 535 286 HIV infected individuals registered in NAP.[16] This provides confidence in the robustness of the algorithm to develop a case report system. Approximately 96% of cases were detected by the algorithm using clinical evidence and only 4% were detected by laboratory information.
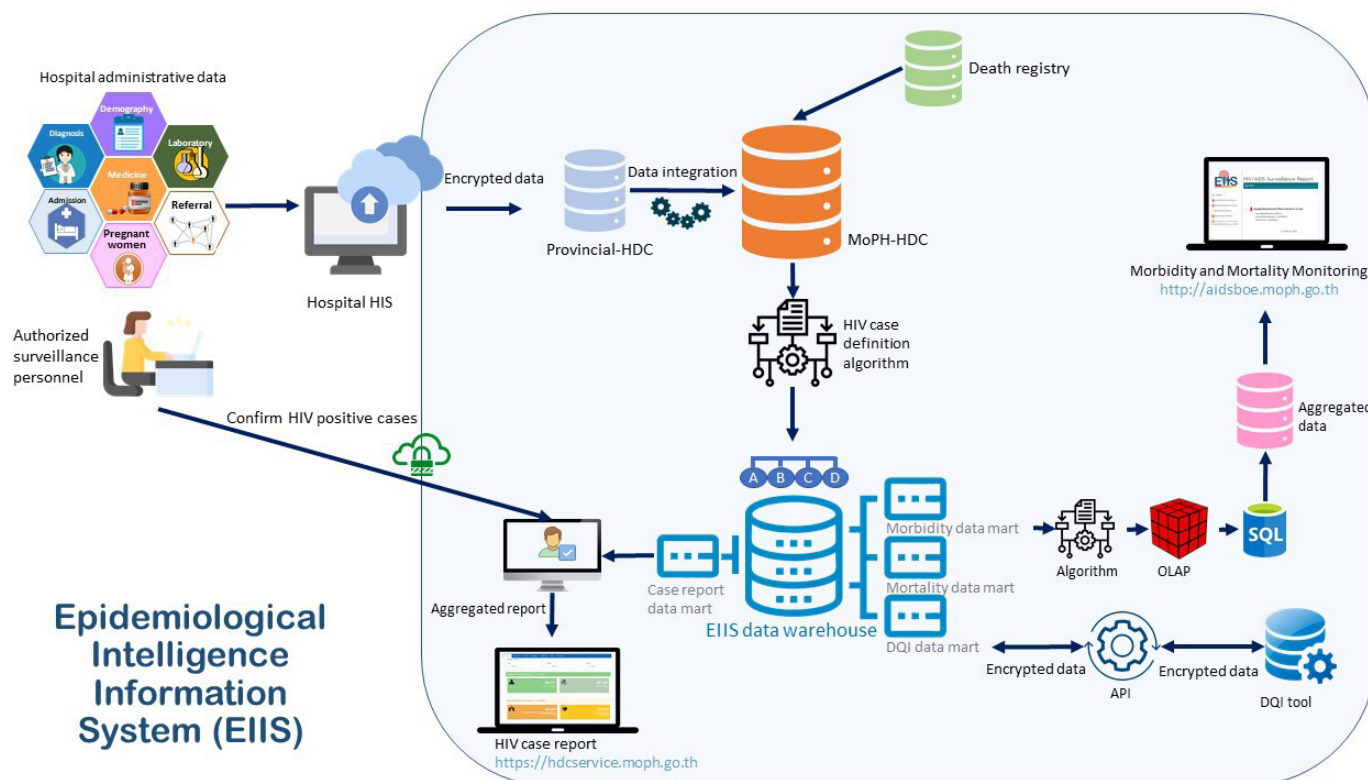
Figure 2 Diagram flow of the EIIS system. API, application programming interface; DQI, data quality improvement; HDC, health data centres; HIS, hospital information system; MoPH, Ministry of Public Health; OLAP, online analytical processing; SQL, structured query language.

The widespread adoption of EMRs for clinical documentation in Thailand has led to unprecedented opportunities for national communicable disease notification and monitoring system. In the past, reporting of notifiable diseases took a considerable amount of time and was performed by aggregating large amount of paper-based information from distant service delivery points up to the programme management level. This process was also prone to errors, loss of information, under-report and duplication when data from all service providers were combined.[17] The time taken in compiling information hindered the ability to respond to HIV situations in a timely manner. Previous studies have revealed considerable under-reporting in the national reporting system, limiting its ability to quantify prevalence or provide reliable estimates of future trends.[18 19]

Our novel algorithmic approach to the identification of HIV infection in the EMR system contributed to the development of an HIV case reporting system in Thailand. In October 2018, Thailand MoPH developed the EIIS to report HIV infection and monitor HIV/AIDS situation at national and subnational levels in Thailand. Figure 2 illustrates the diagram flow of the EIIS system. EMR included demographics, lab, prescription and diagnosis transfer from the HIS to provincial HDC and subsequently to MoPH HDC at least once a month. HDC also received death registry from the Ministry of Interior. EIIS applied the final algorithm to select patients who met the HIV case detection algorithm and transferred these cases to the EIIS data warehouse and case report DataMart for authorised surveillance staff

at each health facility to review and confirm HIV infection status via a secured internet Web browser. According to the Disease Control Law and the MoPH regulation on data security, restricted access to the dataset was applied for only authorised HIV/AIDS disease control staff. Accessing the EIIS dataset required an online registration through the EIIS website and approval from the Department of Disease Control. In addition, EIIS data extraction was modified and further used for other purposes such as HIV morbidity and mortality monitoring system, and data quality improvement programme to track loss to follow-up patients.

Three years after implementation from 1 October 2018 to 30 June 2021, final algorithm as implemented in the EIIS detected and stored 518684 cases in the EIIS data warehouse. Of those, 238067 or 46% of cases were eventually reviewed by authorised healthcare workers at health facilities. Of those reviewed, 0.5% (1240 cases) were misclassified as HIV-infected when they were actually HIV-negative. EIIS can improve the reporting system through early notification and provision of detailed contact information; however, it may be limited in capturing information critical for risk factors and exposures. These data elements are variably documented by clinicians and typically only recorded as free-text rather than structured data. Natural-language processing techniques may improve capacity to report these risks in the future.

There are important limitations to what can be expected of EMR-based reporting systems. The ability of an HIV case detection algorithm relies on quality and completeness

of data. Ongoing efforts to improve EMR data quality are necessary. This system may not be able to capture patients receiving services from private hospitals. However, majority of HIV patients received HIV treatment in public facilities under UC where they submit EMR to MoPH HDC, accounting for 79% of all inpatient beds nationwide. The algorithms detect cases by searching for laboratory tests, prescriptions, and diagnosis codes that in combination are suggestive of HIV infection. The algorithms, therefore, need to be updated when new tests or new drugs are introduced, and new coding systems are implemented (eg, ICD-10). Periodic evaluation and system-wide update to continually calibrate algorithms will ensure efficacy of the HIV reporting system.[20] Only 4% of cases were detected by the laboratory-confirmed evidence and 96% of cases were detected by the algorithm using diagnosis and prescription criteria. The HIV case reporting may be delayed as cases who do not meet laboratory criteria will need to wait for HIV diagnosis and treatment data to determine their HIV status. The time and effort required to complete this process vary widely depending on the quality of records of reportable cases and the availability of clinical staff to review cases flagged by the EIIS system for accuracy. Additional resources and efforts are needed to strengthen the system.

CONCLUSIONS

This national automated case reporting system initiative developed in this study is a model for how EMR can automatically identify HIV-infected subjects. Algorithms incorporating laboratory, standard diagnosis codes and medication prescriptions have the potential to improve completeness and timeliness of HIV reporting system compared with the traditional system. Much effort should be concentrated on improving data quality of national EMR data.

ORCID iD
Suchunya Aungkulanon http://orcid.org/0000-0003-4831-7528

REFERENCES

1 Lazarus R, Klompas M, Campion FX, *et al*. Electronic support for public health: validated case finding and reporting for notifiable diseases using electronic medical data. *J Am Med Inform Assoc* 2009;16:18–24.
2 Jajosky RA, Groseclose SL. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health* 2004;4:29.
3 Thaewnongiew K, Promthet S, Nilvarangkul K, *et al*. The surveillance system in health centers in northeastern Thailand. *Jpn J Infect Dis* 2009;62:444–9.
4 Afzal Z, Engelkes M, Verhamme KMC, *et al*. Automatic generation of case-detection algorithms to identify children with asthma from large electronic health record databases. *Pharmacoepidemiol Drug Saf* 2013;22:826–33.
5 Punpanich W, Ungchusak K, Detels R. Thailand's response to the HIV epidemic: yesterday, today, and tomorrow. *AIDS Educ Prev* 2004;16:119–36.
6 Thai Department of Disease Control. *Communicable disease act, B.E. 2558*. Ministry of Public Health, 2015.
7 Paul DW, Neely NB, Clement M, *et al*. Development and validation of an electronic medical record (EMR)-based computed phenotype of HIV-1 infection. *J Am Med Inform Assoc* 2018;25:150–7.
8 Fasciano NJ, Cherlow AL, Turner BJ. Profile of Medicare beneficiaries with AIDS: application of an AIDS casefinding algorithm. *Health Care Financ Rev* 1998;19:19–38.
9 Goetz MB, Hoang T, Kan VL, *et al*. Development and validation of an algorithm to identify patients newly diagnosed with HIV infection from electronic health records. *AIDS Res Hum Retroviruses* 2014;30:626–33.
10 Fultz SL, Skanderson M, Mole LA, *et al*. Development and verification of a "virtual" cohort using the National VA Health Information System. *Med Care* 2006;44:S25–30.
11 Healthcare resource statistics [Internet], 2021. Available: http://phdb. moph.go.th/
12 MoPH. *Standard operational guideline for data collection and transfereing of data standard format of MoPH, version 2.4*. Nonthaburi: Bureau of Policy and Strategy, MoPH, 2021.
13 Bureau of Epidemiology. Department of Disease Control. Ministry of Public Health. *Guidelines for reporing of notifiable communicable disease under the communicable disease control act B.E. 2558*. Nonthaburi, Thailand, 2015.
14 Thai Health Information Standard Development Center. Thai medicine terminology, 2021. Available: https://www.this.or.th/index.php
15 United Nations Development Programme. *The journey of universal access to antiretroviral treatment in Thailand*. Bangkok, Thailand, 2017.
16 National Health Security Office. NAP Web Report, 2021. Available: http://napdl.nhso.go.th/NAPWebReport/main_rep.jsp
17 Bayer R, Fairchild A. *The role of name-based notification in public health and HIV surveillance*. Geneva, Switzerland: UNAIDS, 2000.
18 Kanlayanaphotporn J, Brady MA, Chantate P, *et al*. Pneumonia surveillance in Thailand: current practice and future needs. *Southeast Asian J Trop Med Public Health* 2004;35:711–6.
19 Wichmann O, Yoon I-K, Vong S, *et al*. Dengue in Thailand and Cambodia: an assessment of the degree of underrecognized disease burden based on reported cases. *PLoS Negl Trop Dis* 2011;5:e996.
20 Gilliam D, Powell J, Haugh E, *et al*. Addressing software security and mitigations in the life cycle. *28th Annual NASA Goddard Software Engineering Workshop*; Proceedings; 3-4 Dec, 2003.