A Digital Case-Finding Algorithm for Diagnosed but Untreated Hepatitis C: A Tool for Increasing Linkage to Treatment and Cure

Brooke Wyatt,1 Ponni V. Perumalswami,1-3 Anna Mageras,1 Mark Miller,1 Alyson Harty,1 Ning Ma,1 Chip A. Bowman4 Francina Collado,1 Jihae Jeon,1 Lismeiry Paulino,1 Amreen Dinani,1 Douglas Dieterich,1 Li Li,1 Maxence Vandromme,1 and Andrea D. Branch1

BACKGROUND AND AIMS: Although chronic HCV infection increases mortality, thousands of patients remain diagnosed-but-untreated (DBU). We aimed to (1) develop a DBU phenotyping algorithm, (2) use it to facilitate case finding and linkage to care, and (3) identify barriers to successful treatment.

APPROACH AND RESULTS: We developed a phenotyping algorithm using Java and SQL and applied it to ~2.5 million EPIC electronic medical records (EMRs; data entered January 2003 to December 2017). Approximately 72,000 EMRs contained an HCV International Classification of Diseases code and/or diagnostic test. The algorithm classified 10,614 cases as DBU (HCV-RNA positive and alive). Its positive and negative predictive values were 88% and 97%, respectively, as determined by manual review of 500 EMRs randomly selected from the ~72,000. Navigators reviewed the charts of 6,187 algorithm-defined DBUs and they attempted to contact potential treatment candidates by phone. By June 2020, 30% (n = 1,862) had completed an HCV-related appointment. Outcomes analysis revealed that DBU patients enrolled in our care coordination program were more likely to complete treatment (72% [n = 219] vs. 54% [n = 256]; P < 0.001) and to have a verified sustained virological response (67% vs. 46%; P < 0.001) than other patients. Forty-eight percent (n = 2,992) of DBU patients could not be reached by phone, which was a major barrier to engagement. Nearly half of these patients had FIB-4 scores ≥ 2.67, indicating significant fibrosis. Multivariable logistic regression showed that DBUs who could not be contacted were less likely to have private insurance than those who could (18% vs. 50%; P < 0.001).

CONCLUSIONS: The digital DBU case-finding algorithm efficiently identified potential HCV treatment candidates, freeing resources for navigation and coordination. The algorithm is portable and accelerated HCV elimination when incorporated in our comprehensive program. (Hepatology 2021;74:2974-2987).

HCV infection remains a major public health threat. Highly effective direct-acting antiviral (DAA) treatments have been available...
since 2014; however, 1 million to 3 million persons in the USA remain infected, and ~50,000 new infections occur every year.\(^{(1)}\) HCV is the leading cause of HCC-related death in the USA and in many other parts of the world.\(^{(2-5)}\) HCV is reported to cause ~20,000 deaths annually in the USA, but the actual death rate could be up to 5-fold higher and exceed 80,000 per year.\(^{(6)}\) The World Health Organization (WHO) and other agencies have set goals to reduce HCV infections and premature deaths\(^{(7,8)}\); however, the USA appears unlikely to meet its WHO 2030 impact targets.\(^{(9)}\)

As a first step toward reducing the HCV-related disease burden, public health agencies have issued a series of screening guidelines. Early advisories focused on patients with specific risk factors.\(^{(4)}\) In 2012, the Centers for Disease Control and Prevention (CDC) expanded screening recommendations to include all baby boomers (i.e., persons born 1945–1965). In early 2020, the CDC and the U.S. Preventive Services Task Force further expanded guidelines to include nearly all adults.\(^{(10,11)}\)

To be effective, screening must lead to treatment, but this often fails to occur. According to the New York City Department of Health and Mental Hygiene (NYC DOHMH), only 62% of residents who had a positive test for HCV RNA in 2015 had received HCV treatment by 2019.\(^{(12)}\) A recent study conducted in Bronx, New York, found that 80% of HCV-RNA–positive samples were collected from patients whose medical record already contained a positive HCV-RNA test,\(^{(13)}\) highlighting the large gap between diagnosis and treatment. Similarly, CDC data indicate that 60% of the 2.4 million persons with chronic HCV in the USA are aware of their positive HCV status, but remain untreated.\(^{(1)}\) Sizable populations of diagnosed-but-untreated (DBU) HCV-infected persons have also been described in Europe, Asia, Central and South America, and Africa.\(^{(5)}\) These findings establish that millions of DBU patients exist and require case finding and outreach.\(^{(3-5,14,15)}\)

Electronic medical record (EMR) usage has expanded in recent years,\(^{(16)}\) providing an opportunity to systematically identify DBU HCV patients on a large scale. We previously developed and used phenotyping algorithms to subtype NAFLD and identify patients with HCC.\(^{(17)}\) Phenotyping algorithms are widely used in research to identify patients with specific diseases\(^{(17-19)}\) and they have the potential to improve health care delivery. The PheKB Web site provides an extensive repository of phenotyping algorithms; however, it does not have an algorithm for HCV,\(^{(20)}\) revealing an unmet need.

To facilitate HCV clinical case finding, we developed a high-quality phenotyping algorithm to automatically identify DBU HCV-RNA–positive patients based on EMR data in EPIC, the most widely used platform in the USA.\(^{(21)}\) The algorithm included elements previously used by the NYC DOHMH\(^{(22)}\) and additional structured and unstructured data fields. Medical record numbers of algorithm-defined HCV treatment candidates were given to patient navigators who manually reviewed charts and reached out to treatment candidates, offering care coordination. We compared outcomes of patients enrolled in our care coordination program to patients who were not enrolled and found a positive association with enrollment. Our project demonstrates the usefulness of computer-assisted HCV case finding. To help others eliminate HCV, we posted the phenotyping algorithm on GitHub.
Materials and Methods

PROJECT DESCRIPTION AND SETTING

This project was carried out to enhance the HCV elimination program in the Mount Sinai Health System. This network provides clinical care at 10 main sites in the greater New York metropolitan area and serves >7 million children and adults. Our program assisted patients receiving liver care at six of these sites (Fig. 1). A computer algorithm was developed to identify living, DBU HCV-infected adults. Four hundred seventy-five patients were enrolled in a nested observational study that evaluated clinical outcomes, including rate of HCV treatment initiation. All study procedures were conducted in compliance with the Helsinki accord; the project was approved by the Institutional Review Board of the Icahn School of Medicine at Mount Sinai (New York, NY). The IRB waived the requirement of informed consent.

OUTCOMES OF INTEREST

The four outcomes of interest are the: (1) positive and negative predictive values (PPV and NPV, respectively) of the phenotyping algorithm; (2) number (percent) of algorithm-defined HCV DBU patients who started HCV treatment before June 2020; (3) factors associated with failure to initiate treatment; and (4) treatment outcomes of patients enrolled in the program’s care coordination program.

THE DIGITAL HCV CASE-FINDING ALGORITHM

The digital phenotyping algorithm uses Java and structured query language (SQL). It was applied to data entered or migrated into our main EMR, EPIC, from January 2003 to December 2017. The algorithm recognizes all U.S. Food and Drug Administration–approved HCV-RNA tests recorded in Mount Sinai’s EPIC and/or Soft Computer Corporation (SCC)
clinical laboratory records, all drugs used to treat HCV (Supporting Table S1), and all International Classification of Diseases (ICD)-9/10 codes for HCV infection (B17.10, B17.11, B18.2, B19.10, B19.20, B19.21, K73.2, K74.60, K74.69, R76.8, and Z86.19). It uses natural language processing, accesses the Mount Sinai death registry, calculates Fibrosis-4 (FIB-4) scores, and infers HCV status using serial alanine aminotransferase (ALT) measurements.

The algorithm selected adults ≥18 years old with an HCV-related entry (HCV-specific ICD-9 or -10 code and/or a clinical laboratory test for HCV antibody and/or RNA). It then selected living DBU HCV-infected patients, defined as either patients whose most recent HCV-RNA test was positive and who had no prescription for an HCV medication, and no record of being deceased, or patients with a positive HCV-RNA test dated after the last prescription for an HCV medication. Because DAAs are highly effective and many patients do not obtain HCV-RNA testing after the end of treatment, the algorithm classified patients who had a prescription for HCV treatment as HCV-RNA negative (cured) unless the EMR had a positive posttreatment RNA test. Of those classified as HCV-RNA positive, it also identified DBU patients at high risk for rapid disease progression, defined as FIB-4 score ≥ 2.67, HIV positive, and/or diabetic; and it distinguished between patients whose most recent HCV-RNA test was before and after January 2014. Patients with ALT ≥ 40 IU were considered to have transaminitis, in accordance with the ALT upper limit of normal defined by others. The algorithm also identified subgroups of patients, such as patients who previously achieved a sustained virological response (SVR) to HCV treatment and patients with positive tests for HCV antibodies and no indication that follow-up HCV-RNA testing was performed.

Because HCV-RNA tests are performed less frequently than ALT measurements, we classified patients as “likely to have an ongoing HCV infection” based on ALT measurements recorded before and after the date of the last recorded HCV-RNA test. Based on our data showing that ALT values decreased by ≥50% in >80% of patients who achieved an SVR, we considered patients to be chronically infected if their last HCV-RNA test was positive and ALT values after that test were within 50% of the ALT values obtained before the positive HCV-RNA test. This method was only applied to patients who had ALT measurements at least 30 days before the RNA test and at least 30 days after the RNA test. If not, they were considered still infected by default.

**EVALUATION OF THE CASE-FINDING ALGORITHM’S PERFORMANCE BY BLINDED CHART REVIEW**

A random number generator was used to select EMRs of 500 patients from among the ~72,000 EMRs with an HCV-related entry. Four trained patient navigators each reviewed 250 EMRs (each record was reviewed twice, with conflicts adjudicated). HCV infection and treatment status were determined by laboratory values and/or physician documentation. Manual review was considered the gold standard. Current HCV infection was indicated by one or more of the following: (1) HCV RNA recorded in the most recent laboratory data; (2) HCV RNA documented in a physician note or media section, with no record of treatment; or (3) HCV infection documented in a physician note. Conversely, cured infection was indicated by a record of an SVR, defined as a negative HCV-RNA test after the end of treatment or HCV cure documented in a physician note. Patients were classified as either (1) being alive and having evidence of current HCV infection or (2) not having evidence of current HCV infection and/or not being alive. The algorithm’s performance was evaluated by comparing its classifications to the gold standard, expressed as percentages. The algorithm’s precision (PPV) equaled the number of living HCV-RNA–positive patients identified by manual review divided by the number the algorithm assigned to this category. The NPV equaled the true negatives (i.e., the number of living patients whose last HCV-RNA test was negative, plus the number of patients with no HCV-RNA test, plus the number of deceased patients according to manual review), divided by the number the algorithm classified as negative (i.e., patients who were deceased and/or had no record of a positive HCV-RNA test or evidence that the final test was negative). Specificity equaled the patients (living or deceased) who were no longer eligible for treatment who were correctly classified by the algorithm divided by all patients no longer eligible for treatment as determined by chart review. Recall (sensitivity) equaled the HCV-RNA–positive
patients correctly classified by the algorithm divided by the HCV-RNA-positive patients identified by chart review. Accuracy equaled the percentage of cases the algorithm classified correctly.

PATIENT NAVIGATION AND CARE COORDINATION

Navigators reviewed the EMRs of 6,187 algorithm-defined DBUs, starting with patients with HCV-RNA tests obtained after 2014 and those with risk factors for liver disease progression (HIV infection and/or diabetes). They attempted to phone treatment candidates, dialing all phone numbers at least three times at varying hours of the day on different days of the week over several weeks, leaving voicemail messages when possible. Navigators did not send short message service communications because recipients may be charged for them. Letters were sent to 619 patients at their last documented address in the EMR after three unsuccessful phoning attempts; however, the yield was too low, and mailing was discontinued. Navigators offered assistance in scheduling an initial appointment with an HCV specialist, using procedures previously shown to be effective, and offered linkage to the project’s care coordinators who served six sites (Fig. 1). Navigators also conducted one round of follow-up outreach to likely treatment candidates who were lost to care after they initially engaged at Mount Sinai sites not served by the project’s care coordinators. All patients enrolled in care coordination had at least one in-person or remote session with a coordinator. Coordinators identified barriers that might impede initiating or completing treatment through a structured psychosocial assessment and open conversations. They tailored a care plan and provided services through SVR at 12 weeks after the end of treatment. Services included scheduling appointments; providing pharmacy and insurance coordination; referring to mental health, substance-use disorder, and social services; and contacting patients at weekly/biweekly intervals to check in, provide treatment adherence and alcohol counseling, as well as tailored health education. Coordinators also conferenced weekly with providers and reported patients’ complaints of side effects or difficulties obtaining and/or taking medication. Coordinators continue to contact patients who have not completed treatment twice-yearly to re-engage them.

STATISTICAL ANALYSIS

Weighted one-sample chi-square tests were used to compare the 6,187 patients whose HCV status has been determined to the entire population of algorithm-defined DBU patients. Bivariate and multivariable logistic regressions were used to identify barriers to starting HCV treatment. Two-sample \( t \) tests were performed to assess differences between patients enrolled or not enrolled in care coordination for continuous variables, chi-square for categorical, using IBM SPSS software (Statistics 25; IBM Corp., Armonk, NY). (32)

Results

DESIGN AND EVALUATION OF THE HCV CASE-FINDING ALGORITHM

A phenotyping algorithm was developed to identify DBU HCV-infected patients according to structured and unstructured data entered into EPIC EMRs between January 2003 and December 2017. It also identified patients who had no record of HCV screening and patients who tested positive for HCV antibodies but had no record of HCV-RNA testing (Fig. 2). Of the ~2.5 million EMRs analyzed, the algorithm classified ~72,000 EMRs as having at least one HCV-related entry (an ICD-9/10 code for HCV and/or a clinical laboratory test result for HCV antibody or RNA), and of these, it classified 10,614 as DBU (Fig. 3). Living patients whose last recorded HCV-RNA test was positive were considered DBU unless they had a prescription for an HCV mediation and no subsequent positive test for HCV RNA. Most algorithm-defined DBU patients had no record of HCV treatment (Fig. 2, orange box, “chronically Infected”), but some failed treatment or became reinfected (Fig. 2, orange box, “positive RNA”).

To evaluate the algorithm’s performance, 500 EMRs were randomly selected from the group of ~72,000 EMRs with an HCV-related entry. According to manual review, the 500 EMRs included 85 living HCV-RNA–positive patients and 425 patients who were deceased and/or not HCV-RNA positive. The algorithm identified 84 EMRs as representing living HCV-RNA–positive patients (DBUs); 74 were confirmed by manual review, and 10 were HCV-RNA
negative at last follow-up according to chart notes (Fig. 4). All 10 misclassifications resulted from entries stored in the media section in a format the algorithm could not interpret. The algorithm’s precision (PPV) was 88% (74 of 84), and its sensitivity was 87% (74 of 85). The algorithm identified 416 patients as not being DBU HCV-infected patients. According to the algorithm, these patients were either deceased, their final HCV-RNA test was negative, they had a prescription for an HCV medication that was not followed by a positive HCV-RNA test, or they did not have any HCV-RNA test result. Manual review confirmed the algorithm’s classification in 405 of 416 cases. The algorithm misclassified 11 HCV-RNA-positive patients as RNA negative. The most common causes of misclassification were chart notes that used idiosyncratic language to report HCV-RNA-positive test results or data from scanned documents stored in the EMR’s media section. The algorithm’s NPV was 97% (405 of 416); its specificity was 98% (405 of 415); and its overall accuracy was 96% (percentage of correctly classified cases; [74 + 405]/500). These results show that the digital phenotyping algorithm is a highly effective case-finding tool that can be used to reduce the human resources needed to find DBU patients based on existing EMR data.

**CHARACTERISTICS OF THE DBU HCV TREATMENT CANDIDATES**

Mean age of the 10,614 algorithm-defined DBU HCV-infected patients was 60 years (SD, 12.6; Table 1). Many had advanced liver disease: 50% had a FIB-4 score ≥ 2.67, indicating that at least half of the population had significant fibrosis. By June 2020, navigators had manually reviewed 6,187 charts and attempted to contact potential treatment candidates by phone. Weighted chi-square tests showed that these 6,187
patients are generally similar to the total group in terms of the percentages of patients in various subgroups with a FIB-4 score ≥ 2.67 and ALT ≥ 40 IU/mL (Table 1). Efforts are underway to contact the remaining 4,427 patients (Fig. 3). The algorithm inferred the HCV-RNA status of these patients, as described in Materials and Methods. Most (77%) are likely to remain HCV infected, given that ALT measurements had not decreased by ≥50% since the last available test for HCV RNA. These patients average 63 years old; 47% have a FIB-4 score ≥ 2.67, and 51% have ALT values ≥ 40 IU/mL, indicating fibrosis and transaminitis.

THE GREATEST BARRIER: OUR INABILITY TO CONTACT DBU HCV PATIENTS BY PHONE OR MAIL

Among the 6,187 algorithm–defined DBU patients, 48% (N = 2,992) could not be reached by phone (Fig. 3). Compared to the others, the patients who could not be reached by phone were younger (mean age, 56 vs. 61 years), less likely to have a FIB-4 score ≥ 2.67 (46% vs. 59%), and less likely to have private insurance (18% vs. 50%; P < 0.001 for all comparisons). As determined by multivariable logistic regression, the factors associated with our inability to contact patients include: no record of an appointment with a liver specialist (OR, 2.5); HIV infection (OR, 2.08); HIV and diabetes (OR, 1.79); sex recorded as other/unknown (OR, 1.54); Medicaid as the only type of insurance (OR, 1.49); homelessness (OR, 1.31); and sex recorded as male (OR, 1.25; Table 2). Letters were sent to 619 of these patients, but only 1 engaged in care as a result, indicating that mailing is not effective in this setting. The great majority (83%) of the patients who could not be contacted by phone are likely to remain HCV infected, given that they had no record of HCV treatment and ALT had not decreased by ≥50%.
ENTRY INTO THE HCV TREATMENT PIPELINE

By June 2020, 31% of the 6,187 algorithm-defined DBU patients had kept at least one HCV-specific appointment (N = 1,862) at our institution or elsewhere or expressed an interest in HCV treatment (N = 39; Fig. 3). We analyzed outcomes in 475 patients who expressed an interest in HCV care and had not started treatment before December 2017. By April 2020, 325 of these patients had initiated HCV treatment. Logistic regression analysis revealed that the odds of starting treatment were inversely related to the number of reasons for treatment delay (P < 0.001; OR, 0.48 per cause of delay; 95% CI, 0.38, 0.61; Table 3). Thirty percent of the 325 patients who started treatment had a positive test for HCV RNA that had been in their EMR >12 months, highlighting the need for proactive programs to identify DBU patients with chronic HCV infection and provide them support services.

Among the 475 patients, 219 enrolled in our care coordination program. Around 30% requested or accepted referrals to support services such as primary care, mental health care, transportation, detoxification/rehabilitation, and housing management. We compared outcomes between these 219 patients and 256 patients who were not enrolled in our program during the same period. The two groups were similar in baseline FIB–4 scores and, among those who started treatment, in the time from initial evaluation to treatment...
### TABLE 1. Characteristics of 10,614 Algorithm-Defined HCV Treatment Candidates Compared to Characteristics of 6,187 Algorithm-Defined HCV Treatment Candidates Whose EMRs Were Manually Reviewed

| Group                        | Total 10,614 | Reviewed 6,187 | PValue† |
|------------------------------|--------------|----------------|---------|
|                              | n            | Age Mean (SD)  | FIB-4* ≥ 2.67, n (%) | FIB-4* Median (IQR) | ALT (IU/L) Median (IQR) | ALT ≥ 40, n (%) | FIB-4* ≥ 2.67, n (%) | FIB-4* Median (IQR) | ALT (IU/L) ≥ 40, n (%) | ALT ≥ 40, n (%) | P FIB-4 ≥ 2.67 | P ALT ≥ 40 |
|                              | n            |               |                       |                  |                        |                   |                     |                       |                            |              |                |              |
| Total                        | 10,614       | 60.2 (12.6)   | 4,196 (50)            | 2.6 (1.5, 6.1)   | 45 (25.69)             | 652 (53)          | 649                 | 79.7 (6.0)            | 381 (73)                  | 4.4 (2.5, 8.8) | 38 (22.66)    | 313 (49)    |
|                              | 6,187        | 60.5 (14.0)   | 649 (49)              | 2.5 (1.4, 5.8)   | 37 (23.62)             | 819 (47)          | 0.78                | 0.27                  |                            | 0.05          | 0.01          |
| Men                          | 6,667        | 59.5 (12.1)   | 2,628 (51)            | 2.7 (1.5, 6.4)   | 48 (30.80)             | 7,185 (51)        | 4,144               | 58.3 (12.2)           | 1,433 (50)                 | 2.7 (1.5, 6.2) | 37 (30.77)    | 2,484 (61) |
|                              | 6,187        | 60.3 (14.0)   | 1,212 (51)            | 2.6 (1.5, 5.9)   | 40 (25.67)             | 2,212 (50)        | 0.78                | 0.27                  |                            | 0.05          | 0.01          |
| Women                        | 3,050        | 61.5 (13.5)   | 1,223 (49)            | 2.6 (1.4, 5.9)   | 40 (25.67)             | 1,496 (50)        | 1.78                | 60.5 (14.0)           | 649 (49)                  | 2.5 (1.4, 5.8) | 37 (23.62)    | 819 (47)    |
|                              | 6,187        | 61.0 (14.0)   | 1,780 (49)            | 2.5 (1.4, 5.8)   | 37 (23.62)             | 819 (47)          | 0.78                | 0.27                  |                            | 0.05          | 0.01          |
| Birth cohort                 |              |                |                       |                  |                        |                   |                     |                       |                            | 0.00          | 0.00          |
| Before 1945                  | 1,238        | 79.7 (5.8)    | 795 (73)              | 4.4 (2.6, 8.9)   | 42 (25.69)             | 652 (53)          | 649                 | 79.7 (6.0)            | 381 (73)                  | 4.4 (2.5, 8.8) | 38 (22.66)    | 313 (49)    |
| 1945-1965                    | 6,515        | 62.1 (5.3)    | 2,810 (52)            | 2.8 (1.7, 6.5)   | 44 (27.74)             | 3,610 (56)        | 3,905               | 61.8 (5.3)            | 1,542 (52)                 | 2.83 (1.7, 6.4) | 43 (27.70)    | 2,109 (55) |
| 1966-1986                    | 1,759        | 43.1 (6.0)    | 246 (24)              | 1.2 (0.8, 2.3)   | 50 (30.80)             | 1,080 (63)        | 1,207               | 42.9 (6.1)            | 157 (24)                  | 1.3 (0.9, 2.5) | 51 (31.83)    | 766 (65)   |
| 1987-2000                    | 207          | 26.8 (2.8)    | 2 (3)                 | 0.6 (0.4, 0.7)   | 61.5 (36.19)           | 140 (70)          | 161                 | 26.7 (2.8)            | 2 (3)                   | 0.6 (0.4, 0.9) | 66 (38.127)   | 113 (73)   |
| Risk factors                 |              |                |                       |                  |                        |                   |                     |                       |                            | 0.00          | 0.00          |
| HIV                          | 1,495        | 56.7 (10.3)   | 572 (47)              | 2.42 (1.5, 5.0)  | 42 (26.99)             | 792 (53)          | 1,226               | 56.7 (10.4)           | 481 (47)                  | 2.5 (1.5, 5.2) | 43 (26.69)    | 659 (54)   |
| Diabetes                     | 1,194        | 65.0 (9.3)    | 561 (50)              | 2.7 (1.6, 5.7)   | 38 (22.63)             | 558 (47)          | 0.39                | 0.77                  |                            | 0.00          | 0.00          |
| HIV/diabetes                 | 256          | 61.1 (7.7)    | 111 (51)              | 2.6 (1.8, 5.2)   | 37 (21.62)             | 120 (47)          | 0.39                | 0.77                  |                            | 0.00          | 0.00          |
| Insurance                    |              |                |                       |                  |                        |                   |                     |                       |                            | 0.00          | 0.00          |
| Medicaid and Medicare        | 206          | 65.2 (9.3)    | 120 (58)              | 3.5 (1.9, 8.0)   | 46 (26.65)             | 111 (54)          | 107                 | 64.5 (8.1)            | 65 (61)                  | 3.8 (2.0, 8.3) | 46 (28.5, 66.5) | 60 (56)   |
| Medicare                     | 1,457        | 69.1 (10.8)   | 836 (57)              | 3.6 (1.9, 7.6)   | 37 (22.64)             | 645 (44)          | 898                 | 68.3 (10.9)           | 515 (57)                  | 3.6 (1.9, 7.7) | 37 (21.62)    | 395 (44)   |
| Medicaid                     | 3,174        | 58.6 (10.8)   | 1,345 (42)            | 2.4 (1.4, 5.5)   | 43 (26.72)             | 1,518 (48)        | 1,739               | 57.6 (11.0)           | 752 (43)                  | 2.5 (1.5, 5.8) | 43 (26.70)    | 875 (50)   |
| Private and/or Medicare      | 5,540        | 58.6 (12.9)   | 1,826 (33)            | 2.6 (1.4, 6.0)   | 48 (30.82)             | 3,087 (56)        | 3,307               | 57.1 (13.1)           | 844 (26)                  | 2.5 (1.4, 5.7) | 47 (30.79)    | 1,893 (57) |
| Uninsured/other              | 238          | 55.8 (12.4)   | 69 (29)               | 1.6 (1.1, 4.4)   | 49 (34.73)             | 123 (52)          | 136                 | 54.9 (12.4)           | 36 (26)                  | 1.6 (1.0, 4.7) | 47 (32.71)    | 80 (59)    |

*The FIB-4 score was calculated using the last laboratory data collected before December 2017.

†Comparing reviewed versus total population proportions with weighted one-sample chi-square test. Bolded values represent statistically significant differences.

Abbreviation: IQR, interquartile range.
|                          | Reached n = 3,185, n (%) | Unreachable n = 3,002, n (%) | Bivariate Logistic Regression | Multivariable Model |
|--------------------------|--------------------------|-------------------------------|-------------------------------|---------------------|
|                         | M (SD)                   | M (SD)                        | P Value | OR     | CI     | P Value | OR     | CI     |
| **Insurance**           |                          |                               |         |   |   |         |         |   |   |
| Private and/or Medicaid and Medicare (ref) | 718 (30.6)               | 387 (23.6)                    | <0.001  | 1.8   | (1.55, 2.11) | <0.001  | 1.49 | (1.25, 1.78) |
| Medicaid                | 880 (37.6)               | 859 (52.3)                    | <0.001  | 0.93  | 0.98   | (0.65, 1.49) | 0.96  | 0.99   |
| Medicaid and Medicare   | 70 (3)                   | 37 (2.3)                      | 0.12    | 0.86  | (0.72, 1.04) | 0.57    | 0.94  |
| Medicare                | 613 (26.2)               | 285 (17.4)                    | 0.001   | 2.21  | (1.55, 3.17) | 0.39    | 1.22  |
| Uninsured/other         | 62 (2.6)                 | 74 (4.5)                      | <0.001  | 0.001 | <0.001 | <0.001  | 1.25 | (1.06, 1.47) |
| **Sex**                 |                          |                               |         |   |   |         |         |   |   |
| Females (ref)           | 1,022 (32)               | 768 (25)                      | <0.001  | 1.4   | (1.25, 1.56) | 0.007   | 1.25 | (1.06, 1.47) |
| Males                   | 2,034 (64)               | 2,110 (70)                    | <0.001  | 1.11  | 1.4    | (1.08, 1.82) | 0.02  | 1.54   | (1.08, 2.21) |
| Other/unknown           | 129 (4)                  | 134 (5)                       | 0.11    | 1.4   | (1.08, 1.82) | 0.02    | 1.54  |
| **HIV/diabetes**        |                          |                               | <0.001  | 1.25  | (1.06, 1.47) | <0.001  | 2.08  | (1.74, 2.48) |
| None (ref)              | 1,986 (62)               | 1,999 (67)                    | <0.001  | 2.08  | (1.74, 2.48) | <0.001  | 2.08  | (1.74, 2.48) |
| HIV                     | 558 (17.5)               | 668 (22)                      | 0.008   | 1.2   | (1.05, 1.40) | <0.001  | 2.08  | (1.74, 2.48) |
| Diabetes                | 531 (17)                 | 220 (7)                       | 0.001   | 0.41  | (0.35, 0.49) | 0.006   | 0.735 | (0.59, 0.92) |
| Both                    | 110 (3.5)                | 115 (4)                       | 0.78    | 1.04  | (0.79, 1.36) | <0.001  | 1.79  | (1.29, 2.49) |
| Homelessness            | 484 (15)                 | 552 (18)                      | 0.001   | 1.25  | (1.10, 1.44) | 0.007   | 1.31  | (1.08, 1.59) |
| Intravenous drug use    | 1,750 (55)               | 1,318 (44)                    | <0.001  | 0.64  | (0.58, 0.71) | 0.11    | 0.87  |
| FIB-4 >2.67             | 1,387 (59)               | 825 (46)                      | <0.001  | 0.75  | (0.67, 0.85) | 0.21    | 0.91  |
| FIB-4                   | 5.1 (5.4)                | 4.3 (4.7)                     | <0.001  | 0.97  | (0.96, 0.98) | <0.001  | 0.97  | (0.96, 0.98) |
| ALT                     | 63.2 (65.6)              | 65.6 (81.8)                   | 0.4     | 1.0   | (1.01, 1.001) | 0.05    | 0.96  |
| Age                     | 61.3 (12.2)              | 56.5 (13.0)                   | <0.001  | 0.97  | (0.96, 0.97) | <0.001  | 0.97  | (0.96, 0.97) |
| No. of phone numbers on file | 2.4 (1.8)               | 2.3 (1.6)                     | 0.001   | 0.95  | (0.92, 0.98) | 0.05    | 0.96  |
| No. of addresses on file | 1.8 (1.3)                | 1.9 (1.4)                     | <0.001  | 1.07  | (1.03, 1.11) | 0.01    | 1.08  | (1.02, 1.15) |
| No liver care at Sinai  | 697 (22)                 | 1,620 (54)                    | <0.001  | 4.2   | (3.75, 4.67) | <0.001  | 2.5   | (2.18, 3.05) |
However, a higher percentage of enrolled patients started treatment (81% vs. 58%; \( P < 0.001 \)), completed treatment (72% vs. 54%; \( P < 0.001 \)), and achieved an SVR, defined as a negative HCV-RNA test \( \geq 4 \) weeks after the end of treatment (67% vs. 46%; \( P < 0.001 \); Table 3). The reasons for not starting and not completing treatment are presented in Supporting Tables S2 and S3, respectively.

**Discussion**

We built an algorithm that uses two standard programming languages, Java and SQL, to identify HCV treatment candidates. It is a useful tool for finding DBU patients with chronic HCV infection and can be widely applied to EPIC EMRs. Based on measured sensitivity, the algorithm identified \(~87\%\) of all living adults whose EMRs contained a most recent HCV-RNA test that was positive. The algorithm reduced the number of charts requiring manual chart review from \(~2.5\) million to 10,614, a 235-fold enrichment, freeing resources for outreach and care coordination. Additional features include its ability to risk-stratify patients based on factors such as persistent ALT elevations, the FIB-4 score, and medical history, including diabetes and HIV infection. The demographics of DBU patients (mean age, 60 years; 60% male) are consistent with USA national data showing that persons born between 1945 and 1969 account for \(72\%\) of chronic HCV infections, and that chronic infection is higher in men.\(^{36,37}\) It is concerning that 50% of DBU patients had a FIB-4 score \( \geq 2.67 \), indicating that many had advanced fibrosis/cirrhosis and are in urgent need of HCV treatment and evaluation for liver cancer surveillance.\(^{33-35}\) Our study revealed the need for additional innovative strategies for engaging DBU patients in HCV care. Only \(~30\%\) received treatment according to our data. The greatest barrier to linking DBU patients to care was our inability to reach them by phone. The NYC DOHMH reported similar findings. Their navigators were only able to reach 42% of the 1,096 patients with viral hepatitis they attempted to contact in 2019.\(^{12}\) We are currently developing information technology (IT)-based approaches to contact persons by direct, secure electronic messages in EPIC (MyChart) and deliver messages to providers with whom these patients are actively engaged.

Fragmentation of the USA health care system is an additional barrier to HCV elimination. IT can reduce this barrier by using surrogate data (i.e., persistence of ALT elevations in patients with previous positive tests for HCV RNA) to bridge gaps in EMRs that may arise because patients obtain care in multiple networks. We created a subprogram that analyzes ALT measurements, which are obtained far more frequently, to manage the frequent absence of the recent HCV-RNA data. Our DBU case-finding algorithm can be adapted for use on other EMRs, providing a tool that can bring a degree of uniformity to HCV case finding.

Our study underscores the importance of active case management. Thirty percent of 325 patients who started treatment in our nested study had a positive HCV-RNA test result that had been in their medical record for >12 months. Many DBU patients had competing medical, financial, psychosocial, and life priorities (such as caregiver obligations) that we attempted to address by providing comprehensive services, including referrals to social services and help coordinating transportation and insurance. Our findings emphasize the value of navigation and care coordination for
HCV treatment, as we and others have demonstrated, and has been demonstrated in other settings. After starting treatment, patients face challenges to completing it that need to be addressed. A multifaceted approach is needed to eliminate HCV within a network like ours (Fig. 5). Complementary initiatives are needed to: (1) contact and engage DBU patients who are identified based on historical data in the EMR; (2) provide a user-friendly digital portal for receiving referrals; and (3) identify previously undiagnosed patients through screening. The CDC’s recommendation to screen nearly all adults was released just as the COVID-19 pandemic was devastating communities around the globe, which likely reduced awareness of the guidelines. We examined screening data collected during the final quarter of 2020 at primary care clinics at Mount Sinai Hospital. Only 13% of previously unscreened adults who attended at least one appointment received HCV screening (unpublished data). The NYC DOHMH estimates that 40% of HCV cases remain undiagnosed in New York City, consistent with CDC data indicating that 39% of infections have not been diagnosed. Automated best practice alerts, smart order sets, and health maintenance messages can promote HCV screening. We are currently building HCV-related directives into our EMR.

**FIG. 5.** Comprehensive HCV elimination across a health care system. Abbreviations: ED, emergency department; OB/GYN, obstetrics and gynecology.

**STRENGTHS AND LIMITATIONS**

A main strength of our study is the portability of the algorithm. The Mount Sinai EPIC EMR includes data from multiple campuses and serves a wide variety of providers who use the EMR in different ways. Data were sourced from SCC software embedded within the EMR, which further diversified the data formats the algorithm is designed to accommodate. As a result, the phenotyping algorithm can be readily adapted for other EPIC EMRs and, potentially, even non-EPIC EMRs. Moreover, the algorithm can be modified to parse incoming data; we are currently using it to identify newly diagnosed patients. In addition, the algorithm can be used to identify patients stalled at various points in the HCV care pipeline, including patients who have not received HCV screening and patients with positive HCV antibody tests and no confirmatory HCV-RNA testing. We deposited our algorithm on GitHub, so that it can be accessed by other health care groups.

Regarding weaknesses, applying the algorithm is low cost and scalable, but it does require the involvement of informatics experts. That said, nearly all USA-based health care practices now use EMRs and have IT personnel, making application feasible, and benefiting health systems that are accountable care organizations.
with full responsibility for the care of their patients. Incomplete medical records were a limitation of this real-world study, as was the inability of navigators to contact many of the treatment candidates by phone. Automated methods to update phone numbers (e.g., when patients make new appointments) may be helpful.

In conclusion, the digital case-finding algorithm is an efficient tool for identifying patients who are HCV-RNA positive and likely treatment candidates and for stratifying based on risk factors for more successful linkage. Once the algorithm is written for a specific EMR, its use is essentially free of cost. Prototypical HCV microelimination projects such as this one are especially important in the USA given the lack of a robust, federally funded plan for the country. (3,37) The approach used here to support an HCV elimination program can be applied to other diseases, such as HBV and metabolic fatty liver disease.

Acknowledgment: Many thanks to our care coordination and navigation team members, Priscilla Agyeman, Lidia Funes, Daryin Hummel, De Shaundra Page-Cook, Glyn Singleton, and Colleen Stapleton.

Author Contributions: Brooke E Wyatt - Equal contributor to conceptualization, data curation and project administration. Lead formal analysis, investigation, methodology, validation, and visualization. Maxence Vandromme - Lead software components. Contributed to formal analysis, investigation, methodology, validation, and visualization. Equal contributor to writing of original draft, review, and editing. Andrea D. Branch - Lead contributor to conceptualization, funding acquisition, project administration and supervision, methodology, original draft writing, review, and editing.

REFERENCES

1) Ryerson AB, Schillie S, Barker LR, Kapronis BA, Wester C. Vital signs: newly reported acute and chronic hepatitis C cases—United States, 2009-2018. Morb Mortal Wkly Rep 2020;69:399-404.
2) Kim HS, El-Serag HB. The epidemiology of hepatocellular carcinoma in the USA. Curr Gastroenterol Rep 2019;21:17.
3) Heffernan A, Cooke GS, Nayagam S, Thurstz M, Hallett TB. Scaling up prevention and treatment towards the elimination of hepatitis C: a global mathematical model. Lancet 2019;393:1319-1329.
4) Stasi C, Silvestri C, Voller F. Update on hepatitis C epidemiology: unaware and untreated infected population could be the key to elimination. SN Compr Clin Med 2020 Oct 18. https://doi.org/10.1007/s42399-020-00058-3. [Epub ahead of print]
5) Saraswat V, Norris S, de Knecht RJ, Sanchez Avila JF, Sondurup M, Zuckerman E, et al. Historical epidemiology of hepatitis C virus (HCV) in select countries—volume 2. J Viral Hepat 2015;22:6-25.
6) Mahajan R, Xing J, Liu SJ, Ly KN, Moorman AC, Rupp L, et al. Mortality among persons in care with hepatitis C virus infection: the Chronic Hepatitis Cohort Study (CHCS), 2006-2010. Clin Infect Dis 2014;58:1055-1061.
7) World Health Organization. Global health sector strategy on viral hepatitis 2016-2021. Towards ending viral hepatitis. Geneva, Switzerland: World Health Organization; 2016.
8) Gaudio A, Gay B, Garmon C, Selick M, Veeland R, Burk K, et al. Localized US efforts to eliminate hepatitis C. Infect Dis Clin North Am 2018;32:293-311.
9) Dore GJ, Martinello M, Alavi M, Grebely J. Global elimination of hepatitis C virus by 2030: why not? Nat Med 2020;26:157-160.
10) Schillie S, Wester C, Osborne M, Wesolowski L, Ryerson AB. CDC recommendations for hepatitis C screening among adults—United States, 2020. MMWR Recomm Rep 2020;69:1-17.
11) U.S. Preventive Task Force. Hepatitis C virus infection in adolescents and adults: screening. Rockville, MD: U.S. Preventive Task Force; 2020.
12) New York City Department of Health and Mental Hygiene (NYC DOHMH). Hepatitis A, B and C in New York City: 2019 Annual Report. Long Island City, NY: NYC DOHMH; 2020.
13) Torian LV, Felsen UR, Xia Q, Larache F, Rude EJ, Rose H, et al. Undiagnosed HIV and HCV infection in a New York City emergency department, 2015. Am J Public Health 2018;108:652-658.
14) Patel AA, Bui A, Prohl E, Bhattacharya D, Wang S, Branch AD, Perumalswami PV. Innovations in Hepatitis C screening and treatment. Hepatol Commun 2020;5:371-386.
15) Yehia BR, Schranz AJ, Umscheid CA, Lo Re V III. The treatment cascade for chronic hepatitis C virus infection in the United States: a systematic review and meta-analysis. PLoS One 2014;9:e101554.
16) Adler-MiLOstein J, Jha AK. HITECH Act drove large gains in hospital electronic health record adoption. Health Aff 2017;36:1416-1422.
markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2009;7:1104-1112.
36) Centers for Disease Control and Prevention. Viral Hepatitis Surveillance—United States, 2018. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2020.
37) Bradley H, Hall EW, Rosenthal EM, Sullivan PS, Ryerson AB, Rosenberg ES. Hepatitis C virus prevalence in 50 US states and DC by sex, birth cohort, and race: 2013-2016. Hepatol Commun 2020;4:355-370.
38) Behrends CN, Eggman AA, Gurkinds S, Bresnanam MF, Fluegge K, Laraque F, et al. A cost reimbursement model for hepatitis C treatment care coordination. J Public Health Manag Pract 2019;25:253-261.
39) Visconti AJ, Sell J, Greenblatt AD. Primary care for persons who inject drugs. Am Fam Physician 2019;99:109-116.
40) McDermott CL, Lockhart CM, Devine B. Outpatient directly observed therapy for hepatitis C among people who use drugs: a systematic review and meta-analysis. J Virus Erad 2018;4:118-122.
41) Paquette CE, Pollini RA. Injection drug use, HIV/HCV, and related services in nonurban areas of the United States: a systematic review. Drug Alcohol Depend 2018;188:239-250.
42) Samuel ST, Martinez AD, Chen Y, Markatou M, Talal AH. Hepatitis C virus knowledge improves hepatitis C virus screening practices among primary care physicians. World J Hepatol 2018;10:319-328.
43) Irvine MK, Chamberlin SA, Robbins RS, Myers JE, Braunstein SL, Mits BS, et al. Improvements in HIV care engagement and viral load suppression following enrollment in a comprehensive HIV care coordination program. Clin Infect Dis 2015;60:298-310.
44) Chubbler NR, Vogel WB, Garel M, Qin H, Koh R, Ryan P. Health services utilization of a care coordination/home-telehealth program for veterans with diabetes: a matched-cohort study. J Ambulatory Care Management 2005;28:230-240.
45) Darvishian M, Wong S, Binka M, Yu A, Ramji A, Yoshida EM, et al. Loss to follow-up: a significant barrier in the treatment cascade with direct-acting therapies. J Viral Hepatitis 2020;27:243-260.
46) Perri-Moore S, Kapssandy S, Doyon K, Hill B, Archer M, Shane-McWhorter L, et al. Automated alerts and reminders targeting patients: a review of the literature. Patient Educ Couns 2016;99:953-959.
47) Al-Hili E, Shankweiler C, Stricklen D, Gibson C, Dunn W. Electronic medical record alert improves HCV testing for baby boomers in primary care setting: adults born during 1945-1965. BMJ Open Qual. 2015;6:e000084.
48) Su J, Lim J. Effectiveness of electronic medical record best practice alert for birth cohort HCV screening in an urban academic primary care clinic. Am J Gastroenterol 2020;115:S547.
49) Carbonneau M, Eboreime EA, Hyde A, Campbell-Scherer D, Faris P, Granlich L, et al. The cirrhosis care Alberta (CCAB) protocol: implementing an evidence-based best practice order set for the management of liver cirrhosis—a hybrid type I effectiveness-implementation trial. BMC Health Serv Res 2020;20:558.
50) Vandromme M. mv5477/hcv-case-finding: HCV case finding. Version 2020. Zenodo 2021 March 5. https://doi.org/10.5281/zenodo.4585942.

Author names in bold designate shared co-first authorship.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.32086/supinfo.