PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Validation of an algorithm using inpatient electronic health records to determine the presence and severity of cirrhosis in patients with hepatocellular carcinoma in England – an observational study
AUTHORS	Driver, Robert; Balachandrakumar, Vinay; Burton, Anya; Shearer, Jessica; Downing, A; Cross, Tim; Morris, Eva; Rowe, Ian

VERSION 1 - REVIEW

REVIEWER	Jennifer Flemming, Assistant Professor Queen's University, Canada
REVIEW RETURNED	13-Feb-2019

GENERAL COMMENTS	I had the opportunity to review the article by Driver et al. titled "Validation of an algorithm using inpatient electronic health records to determine liver disease severity in patients with Hepatocellular carcinoma" currently under consideration for publication in BMJ Open. This was retrospective study of patients with hepatocellular carcinoma from 2007-2016 in the UK that aimed to validate cirrhosis, decompensation, and Baveno stage using chart abstraction as the gold standard using two patient cohorts. The authors conclude that administrative inpatient data can accurately identify the above based on their findings. My comments are outlined below: Major Comments - There is a lack of detail regarding 1) how HCC was diagnosed, 2) all elements of the clinical chart abstraction, and; 3) gold standard definitions of all the conditions being validated (see detailed comments below). Therefore it is difficult to truly evaluate the main findings of the study. - There is no Table 1 describing the clinical characteristics of the two clinical cohorts to determine if they are representative of real world cohorts of patients with HCC. Detailed Comments 1. Introduction - Should also reference LaPointe-Shaw et al. PLoS One 2018 that has validated cirrhosis, HCC, and decompensated cirrhosis in administrative data. 2. Methods - How were the cases of HCC identified from the LTHT? It is unclear what "clinical audit data" refers to - Is there any way to link to outpatient records? - Are the ICD codes from admission diagnosis? Discharge diagnosis? Physician billing? – more details needed - How was HCC diagnosed – imaging? Biopsy? AFP? Combination of them? – need much more detail here. Were these all inpatient diagnosed HCC? What about those diagnosed as outpatient who

	never were admitted to hospital? - It is unclear the rationale for excluding patients seen as tertiary referrals, most patients with HCC would be seen in tertiary care for management, the exclusion of them would be a selection bias as well. - Was a primary chart review done on all patients to define a gold-standard diagnosis of cirrhosis, decompensation, Baveno stage? - There needs to be a clear “gold standard” definition of cirrhosis - There needs to be a clear “gold standard” definition of decompensation - There needs to be a clear “gold standard” definition of all Baveno stages - Three separate people evaluated the abstracted data but who actually abstracted the clinical data? How many people? Were they trained? Was there a standard abstraction form? When was the clinical data abstracted – at the time of HCC diagnosis (that would imply they were all diagnosed as inpatients)? What happened if there was a discrepancy in the assessment by the experts? Did you do an audit to determine the accurateness of data abstraction? - How was the HCC diagnosis date defined? - It is unclear in the methods – were the chart abstractions done at the date of HCC diagnosis? 3. Results - Where is Table 1 outlining the two clinical cohorts? Ie. cause of cirrhosis, MELD, CTP etc. What type of decompensation did they have etc? Stage of HCC at diagnosis? - 90% of patients with HCC have cirrhosis, it is very odd that only 66% of the cohort had HCC, could this be selection bias? What was their underling liver histology? Were they HBV without cirrhosis? This needs to be addressed in the discussion - Table 2 and 3 – where are the 95% confidence intervals for your estimates? 4. Discussion - The limitations regarding external validity need to be addressed, for instance this is not population-based and is limited to patients admitted with cirrhotic complications, excludes those who had tertiary care referral etc. - “robust case note evaluation” it is unclear in the methods that this is the case

REVIEWER	Peter A Richardson Baylor College of Medicine Houston Texas USA
REVIEW RETURNED	01-Mar-2019

GENERAL COMMENTS	[1] I believe that the generalizability of these results are somewhat overstated due to the clearly stated restrictions on the construction of the sample. [2] The work is conducted as a chart-based validation of several groups of ICD10 codes and OPCS4 (the latter possibly mapping onto ICD 10 procedure codes used elsewhere in a straightforward way). The results would then depend upon the validity of the component codes, in particular those for varices and ascites which appear in the Baveno stage definitions. Can the operating characteristics of these as well as ALD and AHF be reconstructed from the records of the chart reviews? [3] In many health systems, PPVs can be expected to be higher for coding at inpatient encounters than outpatient ones. Would that be a

	contributing factor in your results? Moreover, the line between inpatient and outpatient is not uniform over health care systems (or even eras of health care). [4] The title speaks about liver disease severity although the theme is really decompensation. Liver fibrosis (as ascertained by biopsy and estimated by biomarker-based APRI and Fib4, e.g.) is also a factor in liver severity. So, the title should reflect this restriction. [5] The rationale for restricting to patients who have already been diagnosed with HCC is not clearly stated. [6] Otherwise, the study team is to be commended for the careful conduct and clear exposition of this work.

VERSION 1 – AUTHOR RESPONSE

Reviewers' comments	Authors' responses
Reviewer 1	
Should also reference LaPointe-Shaw et al. PLoS One 2018 that has validated cirrhosis, HCC, and decompensated cirrhosis in administrative data.	We agree that this important study should be included. It was published after the initial literature review had been undertaken. Reference number: 11
How were the cases of HCC identified from the LTHT?	Reporting of all cancer cases to the central NHS England registry is mandatory. The diagnosis of HCC is confirmed for all patients in a weekly specialist Hepatobiliary Cancer Multidisciplinary Team (MDT) meetings. Live minutes are taken at these meetings and details collected in the EHR. The cohort was identified from the data submitted to the central registry. Detail added to Methods, p3, para 4
Is there any way to link to outpatient records?	In England, the Hospital Episode Statistics (HES) dataset contains details of all inpatient hospital admissions, including ICD10 and OPCS4 codes. LTHT therefore records these codes (that are also used for billing) and is obliged to submit these to the national HES dataset. The aim of this validation study was to demonstrate that using inpatient records alone is sufficient to identify cirrhosis and grade its severity in patients with HCC. This was done specifically in recognition that the central cancer registry is linked to the HES dataset. This will facilitate studies of outcomes for patients with HCC considering the severity of any accompanying liver disease. Detail added to Introduction, p3, para 2 and 3 and Discussion p9, para 2

outpatient who never were admitted to hospital?	criteria. In cases where a radiological diagnosis was not certain, and where indicated, a targeted liver biopsy was performed. Histological diagnosis was also confirmed in resection specimens and explants following liver transplantation. We included all patients diagnosed with HCC, both as outpatients and inpatients. For those patients diagnosed as an outpatient but never admitted to hospital, we still had access to all of their outpatient records to facilitate case note review, but our algorithm would not be able to interrogate their inpatient record to investigate their cirrhosis characteristics. However, because all HCC treatments (apart from Sorafenib) require a hospital admission, and otherwise patients with cirrhosis require admissions to manage complications we expected the majority of patients to have an inpatient record. Among the 289 patients, 249 (86.2%) had an inpatient record. Among the remaining 40 patients who were never admitted to hospital, 12 had cirrhosis according to casenote review. This suggests that the algorithm missed 12/191 (6.3%) patients with cirrhosis by limiting to inpatient cirrhosis diagnosis codes. Detail added to Methods p3, para 4; Results p7 para 1 and Discussion p10, para 2
Are the ICD codes from admission diagnosis? Discharge diagnosis? Physician billing	The ICD10 codes relate to each completed inpatient hospital episode (including day case procedures such as endoscopy and paracentesis). These codes are used for reimbursement and the codes are also submitted to the HES national dataset. Detail added to Methods p4, para 1
How was HCC diagnosed – imaging? Biopsy? AFP? Combination of them? – need much more detail here. Were these all inpatient diagnosed HCC? What about those diagnosed as	HCC was diagnosed after multidisciplinary review at the Cancer MDT meeting. In most cases, HCC was diagnosed by radiological review of MRI and/ or CT, according to the EASL

It is unclear the rationale for excluding patients seen as tertiary referrals, most patients with HCC would be seen in tertiary care for management, the exclusion of them would be a selection bias as well.	The purpose of this validation study was to assess the performance of an algorithm which relies on inpatient codes related to cirrhosis and its complications. We only had access to the inpatient codes from hospital episodes which occurred at LTHT and RLBUHT (both of which are tertiary centres). We therefore only included those patients who are registered with a Clinical Commissioning Group local to these hospitals, where we would expect them to have received their inpatient cirrhosis care. We agree that this needs clarification: if local patients received tertiary-level care for their HCC at LTHT or RLBUHT, they were not excluded. Detail added to Methods, p3, para 4

Was a primary chart review done on all patients to define a gold-standard diagnosis of cirrhosis, decompensation, Baveno stage?	Yes – a primary chart review was undertaken on all patients.
There needs to be a clear “gold standard” definition of cirrhosis	Patients were defined as having cirrhosis if there was explicit mention of cirrhosis in the clinical record or MDT minutes, evidence of portal hypertension on radiological imaging or endoscopy reports, explicit mention of cirrhosis on liver biopsy or a consistent result on transient elastography Detail added to Methods, p5, para 2
There needs to be a clear “gold standard” definition of decompensation	Decompensation was defined as per the Baveno IV classification as the presence of ascites or bleeding from varices (Baveno 3 or 4). There are limitations of not including hepatic encephalopathy in the clinical definition of decompensation, but we found that this was uncommonly coded in our cohort and so we kept to the same definition of decompensation used by Ratib and colleagues. Details added to Methods p6, para 2 and Discussion p10 para 3
There needs to be a clear “gold standard” definition of all Baveno stages	Baveno stage 1- no varices, no ascites. Baveno stage 2 – non-bleeding varices explicitly mentioned in the clinical records or endoscopy reports, but excluded a report of portal hypertensive gastropathy Baveno stage 3 – explicit mention of ascites in the clinical record, requiring diuretic therapy or paracentesis, but a small volume of ascites only visible on cross-sectional imaging was excluded. Baveno stage 4 – explicit mention of variceal haemorrhage in the clinical record or endoscopy reports Detail added to Methods p6, para 2

Three separate people evaluated the abstracted data but who actually abstracted the clinical data? How many people? Were they trained?	The three researchers are all experienced hepatology fellows, working in this field. They actually abstracted the data from the clinical records themselves. Detail added to Methods, p5, para 2
Was there a standard abstraction form? When was the clinical data abstracted – at the time of HCC diagnosis (that would imply they were all diagnosed as inpatients)? What happened if there was a discrepancy in the assessment by	Yes, there was a standard extraction form comprising the clinical characteristics and blood test results for calculation of MELD/ CP scores. The clinical records were reviewed retrospectively between April -August 2018. The clinical records from the time of HCC

the experts? Did you do an audit to determine the accurateness of data abstraction?	diagnosis were reviewed, however these were more usually outpatient clinics. Blood tests were usually taken at the time of outpatient clinic and these were used. Discrepancies were resolved by consensus view – in these cases this was usually the assessment of the degree of ascites (which can be subjective). The accurateness of data abstraction was determined by review of discrepancies. Detail added to Methods p5, para 2
How was the HCC diagnosis date defined?	This is the date of diagnosis ascribed in the clinical record after confirmation at the weekly Hepatobiliary Cancer MDT meeting. Detail added to Methods p3, para 4
It is unclear in the methods – were the chart abstractions done at the date of HCC diagnosis?	The chart abstractions were done retrospectively between April – August 2018. Detail added to Methods p5, para 2
Where is Table 1 outlining the two clinical cohorts? Ie. cause of cirrhosis, MELD, CTP etc. What type of decompensation did they have etc? Stage of HCC at diagnosis?	This table is now included. Unfortunately we do not have cancer stage at diagnosis. Table 2 included

90% of patients with HCC have cirrhosis, it is very odd that only 66% of the cohort had HCC, could this be selection bias? What was their underling liver histology? Were they HBV without cirrhosis? This needs to be addressed in the discussion	This was also a surprise to us and led to further review of all the “non-cirrhotic” cases. 15 of these patients (15%) had advanced fibrosis in the background liver on histological assessment after biopsy or resection, but this fell short of a diagnosis of cirrhosis. Although LTHT is a tertiary referral centre and a regional centre for liver resection, we attempted to avoid selection bias by only including local patients. Despite this, there appear to be a number of patients who are diagnosed with HCC in the absence of known liver disease. We also note that the “non-cirrhotic” cases were significantly older than the “cirrhotic” cases (baseline Table 2). Unless there was clinical or radiological evidence of cirrhosis, patients were labelled as “non-cirrhotic” according to their clinical records. They may have not been investigated further to establish a diagnosis of cirrhosis if not clinically appropriate. Detail added to Results p 7, para 1 and Discussion p10, para 2

Table 2 and 3 – where are the 95% confidence intervals for your estimates?	Now included.
The limitations regarding external validity need to be addressed, for instance this is not population-based and is limited to patients admitted with cirrhotic complications, excludes those who had tertiary care referral etc.	We agree that these findings are limited by not being population-based, but the results of the external validation using another UK centre are consistent. The reliance on inpatient codes within the HES dataset is a limitation. Nevertheless, this is true of other studies that utilise the HES dataset when linked to the national cancer registry and these have led to impactful analyses with generalizable results. Detail added to Discussion, p10, para 1 Added references 22, 23

Reviewer 2	
I believe that the generalizability of these results are somewhat overstated due to the clearly stated restrictions on the construction of the sample.	We acknowledge these concerns over the construction of the sample, and we have therefore clarified the necessity for including only inpatient codes in the Introduction. The external validation using another UK centre has produced consistent results which demonstrates the generalizability within England. This study demonstrates the potential of the linked Inpatient HES dataset to utilise meaningful analysis of the national cancer registry, including adjustment for cirrhosis severity. Similar to previous studies using these linked datasets, we expect those analyses be generalizable to health systems outside the UK. Detail added to Introduction p3, para 2 & 3, and Discussion p9, para 4
The work is conducted as a chart-based validation of several groups of ICD10 codes and OPCS4 (the latter possibly mapping onto ICD 10 procedure codes used elsewhere in a straightforward way). The results would then depend upon the validity of the component codes, in particular those for varices and ascites which appear in the Baveno stage definitions. Can the operating characteristics of these as well as ALD and AHF be reconstructed from the records of the chart reviews?	The new Supplementary Table 3 describes the operating characteristics for detection of varices and bleeding varices. Detection of variceal bleeding based on ICD10 I85.0 and I98.3 alone (Algorithm A) is inferior to using the additional OPCS4 variceal treatment codes used in algorithm B and C. Although algorithm B is more sensitive for detecting variceal bleeding than algorithm C, it incorrectly classifies all patients who have had treatment as having bled (including prophylactic variceal banding) and the PPV falls to 54%. The overall agreement with Baveno stage and decompensation status is optimised using algorithm C. Supplementary Table 4 describes the operating characteristics for detection of ascites using ICD10 R18.X alone, and with the addition of the OPCS4 codes for paracentesis. Including the treatment code improves the sensitivity for detecting ascites. Detail added to Results p8, para 1

In many health systems, PPVs can be expected to be higher for coding at inpatient encounters than outpatient ones. Would that be a contributing factor in your results? Moreover, the line between inpatient and outpatient is not uniform over health care systems (or even eras of health care).	We agree with this observation. The PPV is high in our cohort in part due to the high pre-test probability of cirrhosis in an HCC population. When we tested previous algorithms in our cohort, the operating characteristics outperformed their derivation cohorts.
	We also agree that the distinction between inpatient and outpatient is not uniform across healthcare systems. The codes utilised in this study include those generated from inpatient hospital episodes and day case treatments, but it does not rely on coding of outpatient clinic consultations, which lack ICD10 codes in UK datasets for population studies. Detail added to Discussion, p9, para 2
The title speaks about liver disease severity although the theme is really decompensation. Liver fibrosis (as ascertained by biopsy and estimated by biomarker-based APRI and Fib4, e.g.) is also a factor in liver severity. So, the title should reflect this restriction.	We agree with this observation, particularly with reference to fibrosis which is not captured by this approach. We have therefore changed the title to "Validation of an algorithm using inpatient electronic health records to determine the presence and severity of cirrhosis in patients with hepatocellular carcinoma in England – an observational study" Title changed
The rationale for restricting to patients who have already been diagnosed with HCC is not clearly stated.	We agree that this important observation has not been clearly addressed in the Introduction. The overarching aim is to exploit the robust patient-level data contained within the national cancer registry for population studies in HCC. Since the registry lacks data about underlying cirrhosis, the purpose of this validation study was to assess the performance of an algorithm which relies on inpatient codes alone to identify cirrhosis and its severity, because the national HES dataset does not contain outpatient ICD10 codes. Detail added to Introduction p3, para 2 and 3

During the preparation of the manuscript for re-submission it was noted that the quoted performance characteristics for detection of decompensation had been made using patients identified with cirrhosis using the algorithm rather than case note review. This has been changed in the text in the Abstract and Results (p7, para 2 and p8, para 2).

VERSION 2 – REVIEW

REVIEWER	Jennifer Flemming, Assistant Professor Queen's University, Canada
REVIEW RETURNED	01-Apr-2019

GENERAL COMMENTS	The authors have addressed all concerns.

REVIEWER	Peter A Richardson Baylor College of Medicine Houston Texas USA
REVIEW RETURNED	16-Apr-2019

GENERAL COMMENTS	The authors' responses were forthright and clear. The revised manuscript is well written and I think is a useful piece to add to the edifice of work in the epidemiology of liver disease in the post HCC diagnosis period.