

Analytical validation and algorithm improvement of HepatoPredict kit to assess hepatocellular carcinoma prognosis before a liver transplantation

Maria Gonçalves-Reis^a, Daniela Proença^a, Laura P. Frazão^a, João L. Neto^a,
Sílvia Silva^b, Hugo Pinto-Marques^{b,c}, José B. Pereira-Leal^a, Joana Cardoso^{a,*}

^a Ophiomics – Precision Medicine, Lisbon, Portugal

^b Hepato-Biliary-Pancreatic and Transplantation Centre, Curry Cabral Hospital, Centro Hospitalar Universitário de Lisboa Central, Lisbon, Portugal

^c Chronic Diseases Research Center (CEDOC), NOVA Medical School, Universidade NOVA de Lisboa (NMS/UNL), Lisbon, Portugal

ARTICLE INFO

Keywords:

HepatoPredict
Hepatocellular carcinoma
Liver transplantation
Analytical validation
Multi-target genomic assay
Prognostic

ABSTRACT

Objectives: To verify the analytical performance of the HepatoPredict kit, a novel tool developed to stratify Hepatocellular Carcinoma (HCC) patients according to their risk of relapse after a Liver Transplantation (LT).

Methods: The HepatoPredict tool combines clinical variables and a gene expression signature in an ensemble of machine-learning algorithms to forecast the benefit of a LT in HCC patients. To ensure the accuracy and reliability of this method, extensive analytical validation was conducted to verify its specificity and robustness. The experiments were designed following the guidelines for multi-target genomic assays such as ISO201395-2019, MIQE, CLSI-MM16, CLSI-MM17, and CLSI-EP17-A. The validation process included reproducibility between operators and between RNA extractions and RT-qPCR runs, and interference of input RNA levels or varying reagent levels. A recently retrained version of the HepatoPredict algorithms was also tested.

Results: The validation process demonstrated that the HepatoPredict kit met the required standards for robustness ($p > 0.05$), analytical specificity (inclusivity of 95 %), and sensitivity (LoB, LoD, linear range, and amplification efficiency between 90 and 110 %). The operator, equipment, input RNA, and reagents used had no significant effect on the HepatoPredict results. Additionally, the testing of a recently retrained version of the HepatoPredict algorithm, showed that this new version further improved the accuracy of the kit and performed better than existing clinical criteria in accurately identifying HCC patients who are more likely to benefit LT.

Conclusions: Even with the introduced variations in molecular and clinical variables, the HepatoPredict kit's prognostic information remains consistent. It can accurately identify HCC patients who are more likely to benefit from a LT. Its robust performance also confirms that it can be easily integrated into standard diagnostic laboratories.

* Corresponding author. Ophiomics – Precision Medicine, Pólo Tecnológico de Lisboa, Rua António Champalimaud, Lote 1 (Sala 14), 1600-514, Lisbon, Portugal.

E-mail address: jvaz@ophiomics.com (J. Cardoso).

<https://doi.org/10.1016/j.plabm.2024.e00365>

Received 1 August 2023; Received in revised form 22 January 2024; Accepted 23 January 2024

Available online 5 February 2024

2352-5517/© 2024 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Background

Primary liver cancer is a significant global health issue, ranking as the 6th most diagnosed cancer and being the third leading cause of cancer death worldwide [1]. Hepatocellular carcinoma (HCC) comprises 75–85 % of primary liver cancer cases and is linked to chronic hepatitis B virus (HBV) or hepatitis C virus (HCV) infection, heavy alcohol consumption, smoking, type 2 diabetes, excess body weight, and the consumption of aflatoxin-contaminated foods [1]. Around 30 % of HCC cases are considered for treatment with curative intent [2] which involves liver transplantation (LT) or surgical resection [2,3]. Liver transplantation is preferred over surgical resection because it treats the HCC and the underlying cirrhosis, reducing the risk of death in the first two years following diagnosis [3]. However, the shortage of liver donors has led to the development of various criteria for identifying HCC patients most likely to benefit from LT [4–13] mainly based on tumor burden and protein biomarkers such as alpha-fetoprotein (AFP) and des-carboxyprothrombin (DCP). Despite all the LT criteria, their limitations are currently under discussion, mainly because they exclude patients with an underlying good prognosis who can benefit from LT and include bad prognosis patients who will not benefit from the surgery [14,15].

The HepatoPredict kit was designed to predict which HCC patients will benefit from LT by determining their prognosis. This is accomplished by combining three clinical variables (tumor number, size of the largest nodule, and total tumor volume) with a gene expression signature that includes *DPT*, *CLU*, *CAPNS1*, and *SPRY2* genes, using a proprietary algorithm. The HepatoPredict kit can extract RNA from formalin-fixed paraffin-embedded (FFPE) HCC samples and perform gene expression analysis using real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) technology. The RT-qPCR results are then combined with the clinical data using a machine learning algorithm that returns three different values concerning the predictive value (Class I and II) or its absence (Class 0) [16].

In a previous study [16], we demonstrated that in the selection of patients suitable for LT HepatoPredict outperforms the Milan [4], University of California San Francisco (UCSF) [6], Up-to-seven [11], alpha fetoprotein (AFP) [7], Metroticket 2.0⁵, Total Tumor Volume (TTV) [13], and TTV AFP [12] criteria in a retrospective clinical validation cohort of patients diagnosed with HCC [16].

In addition to its clinical utility, a prognostic test used in diagnostics must also be able to produce accurate, reliable, and specific data under standard laboratory conditions (analytical validation). This study aims to provide evidence of the analytical validation of the HepatoPredict kit through comprehensive technical analysis. The validation process also involved re-training the HepatoPredict algorithm using a retrospective cohort of 162 patients diagnosed with HCC and who underwent LT.

2. Methods

Samples: In this study, HCC samples preserved as FFPE tissue were used. FFPE HCC samples were acquired from four different suppliers: Biobank IRBLleida (PT20/00021), integrated in the Spanish National Biobanks Network and Xarxa de Bancs de Tumors de Catalunya (XBTC) sponsored by Pla Director d'Oncologia de Catalunya; Biobank ISABIAL, integrated in the Spanish National Biobanks Network and in the Valencian Biobanking Network; and biorepositories from Amsbio (US) and Biotech (US). All samples were processed following standard operating procedures with the appropriate approval of the Ethical and Scientific Committees. Moreover, clinical samples from a retrospective clinical study approved by the ethics authorities and taking place in the Curry Cabral Hospital (Lisbon, Portugal), were also used. All HCC FFPE samples were acquired either sectioned with 3–5 µm thickness or as paraffin blocks that were then cut in 3–5 µm thick slices using a microtome (Leica SM2010R Sliding Microtome, Leica Biosystems) and mounted on a glass slide.

Histopathologic analysis: Prior to RNA extraction, HCC FFPE samples were analyzed by certified pathologists using a hematoxylin and eosin (H&E) stained tumor section. FFPE HCC slides (3 µm thick) were first deparaffinized and stained using Harris Hematoxylin solution (#3801561 E, Leica Biosystems, Richmond, USA). The slide was then counterstained with Eosin Y solution (#2801601, Leica Biosystems, Richmond, USA). Finally, slides were dehydrated in increasing alcohol concentrations, cleared in xylene (#28973, VWR, Alfragide, Portugal), and mounted using a xylene-based mounting medium (#107961, Merck, Darmstadt, Germany). After H&E staining, slides were observed under an optical upright microscope (Panther L, #1100104600142, Motic®).

RNA extraction: For RNA extraction, an HCC area mimicking a needle biopsy was delimited in two sequential 5 µm slides. Samples were initially deparaffinized and the RNA was extracted using the RNeasy FFPE Kit (#73504, Qiagen, Hilden, Germany), according to the manufacturer's instructions with two exceptions: proteinase K cell lysis and final elution volume.

DNA extraction: DNA was extracted from HCC FFPE samples using the QIAamp DNA FFPE Tissue Kit (#56404, Qiagen, Hilden, Germany) in accordance with the manufacturer's instructions.

RNA extraction method validation: For the validation of the RNA extraction method, 87 FFPE HCC samples (from the four different suppliers) were used. Each sample was tested in duplicate and by two different operators on different equipment. Moreover, mirror sections of each sample were used to reduce the variability between operators. Immediately after extraction, RNA was stored at –20 °C until further usage or used straight away in RT-qPCR reactions (conditions described below) targeting *RPL13A* (reference gene) and *Chr3* (genomic DNA control) to analyze the integrity [17] and the purity of the RNA samples, respectively.

RT-qPCR reactions: 1-step RT-qPCR reactions were performed as previously described [16]. The QuantStudio Design & Analysis Software v1.5.1 software was used for data acquisition and analysis. For gene expression normalization, the geometric mean of the cycle threshold (Cq) of the reference genes (*RPL13A*, *GAPDH*, and *TBP*) was subtracted from the Cq values of the genes of interest (*DPT*, *CLU*, *CAPNS1*, and *SPRY2*).

Primer Specificity: RT-qPCR products were sequenced via Sanger sequencing outsourced to Eurofins (<https://eurofinsgenomics.eu/en/custom-dna-sequencing/gatc-services/supremetube/>). In total, 16 different solutions (forward and reverse for 8 targets) were

sent to Eurofins. Regarding RT-qPCR products, a 2-step RT-qPCR reaction was performed using the SuperScript™ VILO™ cDNA Synthesis Kit (#11754050, Thermo Fisher Scientific, Bleiswijk, Netherlands) to synthesize cDNA and the Invitrogen™ Platinum™ SuperFi™ PCR Master Mix with the SuperFi™ GC Enhancer (#12358010, Thermo Fisher Scientific, Bleiswijk, Netherlands) in qPCR. An RNA pool (composed of 8 different FFPE HCC samples) was used as template. The size of each RT-qPCR product was assessed by electrophoresis in a 4 % agarose gel (#G401004, Thermo Fisher Scientific, Bleiswijk, Netherlands), using a DNA ladder (#10488096, Thermo Fisher Scientific, Vilnius, Lithuania) and nuclease-free water (#129114, Qiagen, Hilden, Germany) as a negative control in an electrophoresis system (#G8300, Thermo Fisher Scientific, Vilnius, Lithuania).

RT-qPCR inclusivity: Different FFPE HCC samples were used as templates in RT-qPCR reactions targeting all the genes included in HepatoPredict kit. Each RT-qPCR reaction was performed in duplicate and by two different operators. For each sample, each operator used the same batch of extracted RNA or DNA. Both nucleic acid extraction and RT-qPCR reactions were performed as described above.

Limit of Detection (LoD) determination: For LoD determination, a pool composed of 8 FFPE HCC samples was used. The samples composing the pool reflected high and low expression levels of each target and were associated with a bad prognosis (recurrence, $n = 4$) and a good prognosis (no recurrence, $n = 4$). Both pools, of DNA and RNA, were created using the same samples. DNA pool was directly used for serial dilutions (at least 11 per target) and the RNA pool was diluted 1:4 to create the starting sample for the serial dilutions. In total, 21 replicates were done for each dilution (triplicates in each of the 7 RT-qPCR reactions), per lot number of reagents, on three different days (2–3 RT-qPCR reactions per day). For each target, all reactions were performed by the same operator with the same equipment. RT-qPCR reactions using reagents from different lots were analyzed separately. Data was analyzed in accordance with the Probit model, which implied the creation of a regression representing the probability vs \log_2 dilution for each target assuring at least 3 dilutions with hit rates within 0.10–0.90 and at least one exceeding 0.95. Moreover, to minimize the influence of the model limit ranges of probability, dilutions with a 100 % fail or success rate were included in each analysis. LoDs were independently calculated for each lot and the maximum LoD (concentration) was taken as the reported value for the measurement procedure. To determine the Cq value associated with the LoD, a linear regression was applied between the Cq values and the \log_2 (dilution factor). All the \log_2 (dilution factor) until the one immediately after the LoD were considered. The values outside the confidence interval (CI) at 99 % were considered outliers and were removed. Linear, quadratic, and cubic polynomial functions were fitted to the Cq values using \log_2 dilution values. If none of the non-linear coefficients was different from zero, the target was considered linear (*GAPDH*, *TBP* and *Chr3*). Otherwise, the absolute difference between the model that best fits the data (smallest mean squared error) and the linear model was calculated. When the difference was less than 1 Cq value, the target was considered linear (*RPL13A*, *DPT*, *CAPNS1*, *CLU* and *SPRY2*).

Linearity: The linear range of each target included in the HepatoPredict kit was determined for RT-qPCR reactions using FFPE HCC samples (previously used for LoD determination) and reference materials (#636690, Takara, Saint Germain en Laye, France) to cover a broad range of nucleic acids concentrations in linearity determination. For that, seven serial dilutions of reference RNA were used with 3 replicates per dilution and at least eleven dilutions of nucleic acids pools, obtained from HCC FFPE samples, were used with 7 replicates per dilution. Finally, the Cq values and the dilution factors were plotted in a base 2 logarithmic graph and R^2 (>0.90) was calculated for all targets.

Amplification efficiency: The reaction efficiency was calculated for each target included in the HepatoPredict kit. It was determined from the slope of the log-linear portion of each target curve: amplification efficiency = $(2^{-1/\text{slope}} - 1) \times 100$.

Robustness of RT-qPCR reactions: Plackett and Burman tables [18] were used to design the robustness assay: alterations in the concentrations (± 30 %) of the master mix (#A15300, Thermo Fisher Scientific, Bleiswijk, Germany), primers and probes were implemented as well as different final reaction volumes (± 5 %) and annealing temperatures (± 1 °C), as demonstrated in Supplementary File 1. Two independent assays per each target were performed using the same sample pool (see LoD) in triplicate and all reactions were performed using sample concentration near the LoD (RNA pool serial dilution 2^{-2} and DNA pool without further dilutions). Three standard conditions were incorporated in the assay for data analysis: standard (STD) (no changes), STD1 (-1 °C annealing temperature), and STD2 ($+1$ °C annealing temperature).

Determination of Cq values below LoD and within the linear range for each target: Serial dilutions of the reference RNA (#636690, Takara, Saint Germain en Laye, France) were used as templates for RT-qPCR reactions targeting all genes included in the HepatoPredict kit. The Cq values above the LoD and outside the linear range for each target were identified and the maximum acceptable Cq value for each target was determined.

Precision studies: The conditions under which repeated measurements were made determine the type of precision being analyzed – reproducibility (daily, lot-to-lot, operator, and inter-assay) and repeatability. For the daily reproducibility, for the same sample, assays were performed by the same operator, using the same sample and kit's lot on 4 different days. Regarding lot-to-lot reproducibility, the same sample was analyzed by the same operator using kits from three different lots. Finally, operator reproducibility was studied by using the same sample with HepatoPredict kits from the same lot but performed by three different operators. For each condition, two HepatoPredict kits were used (two independent assays). Repeatability was measured considering the triplicates of each HepatoPredict kit run. In total, 3 different HCC FFPE samples were studied, thus, 48 HepatoPredict kits were used (16 kits/sample). To further assess inter-assay reproducibility, 15 additional samples were tested in duplicate by different operators, using different lots of the HepatoPredict kit, and on different days (total $n = 18$).

HepatoPredict algorithm training: A dataset with 162 patients diagnosed with HCC and submitted to liver transplant, of which 131 were previously studied [19], was used (Supplementary File 2). Concerning the dataset, different models were tested, such as Naive Bayes, support-vector machine (SVM) with different kernel functions, and Extreme Gradient Booster (XGBoost). Moreover, synthetic minority oversampling technique (SMOTE) was also used for data imbalances. Python 3.8 was used with scikit-learn 1.0.2 (<https://scikit-learn.org/stable/>), XGBoost 1.6.1 (<https://xgboost.readthedocs.io/en/stable>), imbalance-learn (<https://imbalanced-learn.org/>

stable/), and Optuna 2.10.0 (<https://optuna.org/>). Furthermore, the leave-one-out cross-validation (LOOCV) approach was used for algorithm training. Each model was fed with 4 molecular (*DPT*, *CLU*, *CAPNS1*, and *SPRY2* gene expression) and 3 clinical variables (tumor number, largest tumor size and total tumor volume). As previously described [16], the algorithm was developed as a two-level predictor.

HepatoPredict algorithm univariate analysis: The univariate analysis of the HepatoPredict algorithm consisted in calculating the error (i.e., counting each time the algorithm would fail the correct prognosis classification) when altering the Cq mean values of each gene (prior normalization) and varying the normalized Cq values of *DPT*, *CLU*, *SPRY2* and *CAPNS1* and the clinical variables (tumor number, diameter of the largest tumor, and total tumor volume). Thus, Cq mean values of the genes of interest (*DPT*, *CLU*, *SPRY2* and *CAPNS1*) were replaced by 40 Cq (the maximum number of cycles allowed) and their respective LoD and the Cq means of the reference genes (*RPL13A*, *GAPDH*, and *TBP*) were removed and replaced by their respective LoD. Regarding the variations of the normalized Cq values (for *DPT*, *CLU*, *SPRY2* and *CAPNS1* genes), variations of 0.1 Cq were performed. Alterations in clinical variables included the variations in the tumor number (1–2 units), and in the diameter of the largest tumor (cm) and the total tumor volume (cm³) by 2 %.

HepatoPredict algorithm multivariate analysis: The multivariate analysis of the HepatoPredict algorithm consisted in calculating the error associated with the alteration of more than one variable at a time. Thus, a range for each variable variation was defined (based on algorithm univariate analysis) and random combinations of 2, 3, 4, 5, 6 and 7 altered variables were tested. All possible combinations of variables were performed and for each combination the assay was repeated 10,000 times with random variable alterations within the defined range. Finally, two types of errors were calculated: error type A (between Class I and Class II) and error type B (between Class I or II and Class 0).

Statistical analysis: Statistical analysis was performed using the R language for Statistical Computing (v 4.1.1) and GraphPad Prism 7 (GraphPad Software, Inc. 2016). For RNA extraction validation, as the data followed a normal distribution, the Paired Student t-test was applied. For robustness assay, Dunn's multiple comparisons test was applied. Regarding precision, due to data size, a non-parametric test (Friedman test) was used. A $p < 0.05$ was considered statistically significant.

3. Results

3.1. Validation of RNA extraction method

For the RNA extraction from HCC FFPE tissues, the RNeasy FFPE kit was used. This RNA extracted method was validated by using 87 HCC FFPE samples handled by two different operators. After extraction, RNA was used as a template in RT-qPCR reactions targeting the *RPL13A* gene and a DNA-specific target (*Chr3*), assuring both the integrity and the purity of the extracted RNA. Each sample was analyzed in duplicate by each operator. Regarding *RPL13A* expression, no significant differences were observed between operators for each sample ($p = 0.27$, Fig. 1A) and the mean standard deviation (SD) between samples tested by the two operators was 0.47 Cq (Fig. 1D). Moreover, no statistically significant differences were observed between each operator's duplicates ($p = 0.99$ for operator 1 and $p = 0.13$ for operator 2, Figs. 1B and 2C respectively). The mean SD between duplicates of each sample for operator 1 was 0.12 Cq while it was 0.18 Cq for operator 2 (Fig. 1D). Regarding *Chr3*, residual genomic DNA (gDNA) contamination (Cq mean > 34) was identified in 8 samples (9.19 %) handled by operator 2 and in one sample (1.15 %) handled by operator 1 (data not shown).

3.2. Analytical specificity – primer specificity and RT-qPCR inclusivity

To demonstrate primers' uniqueness for each target, primers and RT-qPCR products were sequenced and the specificity of each

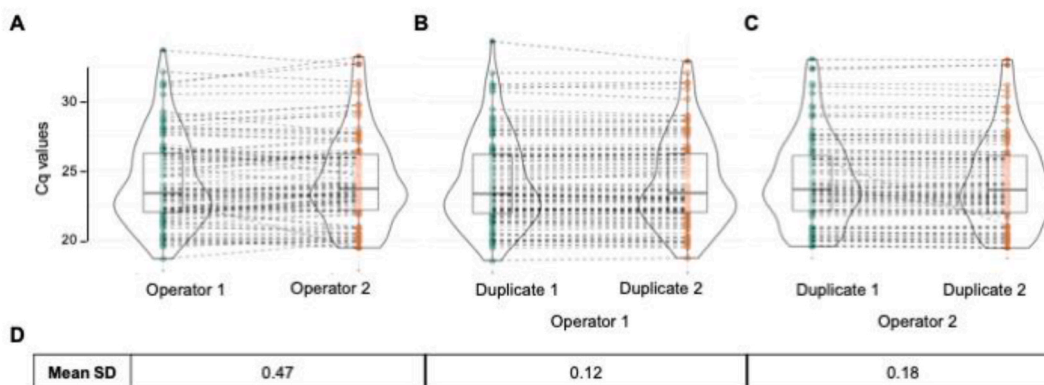


Fig. 1. RNA extraction validation using *RPL13A* gene expression. The RNA extraction method was validated by two different operators ($p = 0.27$, Paired Student's t-test) (A). Moreover, each operator performed the assays in duplicate: operator 1 ($p = 0.99$, Paired Student's t-test) (B) and operator 2 ($p = 0.13$, Paired Student's t-test) (C). The mean SD between operators and within operators is also represented (D). For each violin plot (A, B, and C), dots represent Cq mean values (A) and Cq values (B and C) of *RPL13A* gene. Dashed lines represent the correspondence of samples between groups.

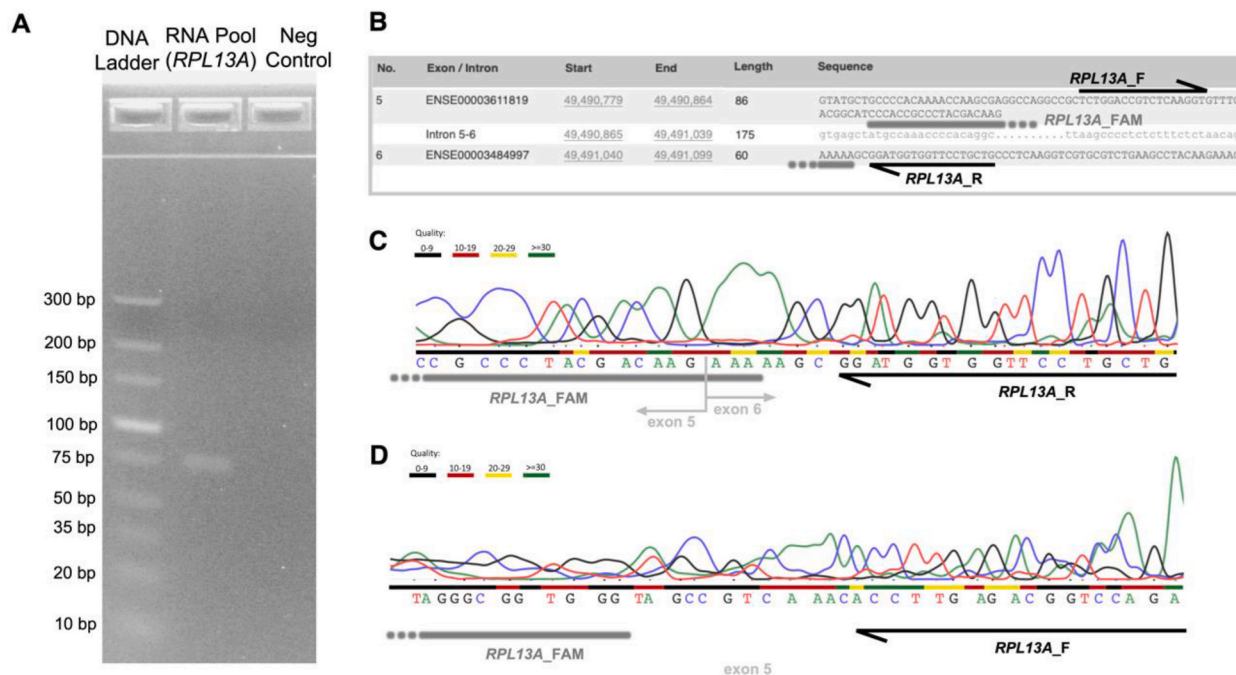


Fig. 2. Primer Specificity (for *RPL13A* as an illustrative example). After a RT-qPCR reaction targeting the *RPL13A* gene, the size of the amplicon was assessed by electrophoresis in an agarose gel (~75 bp), using a DNA ladder and a negative control (nuclease-free water) (A). ENSEMBL canonical transcript sequence (ENST00000391857.9, RefSeq NM_012423), represented from 5' to 3' (B). Alignment of the *RPL13A* forward (*RPL13A_F*) and reverse (*RPL13A_R*) primers and probe (*RPL13A_FAM*) with the Sanger sequencing electropherogram results for the *RPL13A* amplicon in forward (5' to 3') (C) and reverse (3' to 5') (D) directions. The presented image is cropped. The full-size original image can be found in [Supplementary File 4](#).

primer pair was confirmed by aligning the Sanger sequencing electropherograms from each primer with the respective PCR amplicon, as suggested by ISO 20395:2019 and MIQE guidelines [20,21]. Before sequencing, the amplicon size was verified by electrophoresis in an agarose gel (Fig. 2A, *RPL13A* depicted as an example). For the *RPL13A* gene, the electrophoresis band corresponded to 75 bp (Fig. 2A), in accordance with the expected amplicon size [16]. All the amplified amplicons (for the remaining HepatoPredict targets) corresponded to the expected size [16] and no extra bands of unspecific PCR products were observed (data not shown) confirming the specificity of the primer pairs for the desired target. The sequences of all primers and probes were successfully aligned for all targets included in the HepatoPredict (data not shown) in the DNA sequence displayed in the Sanger electropherogram of the respective PCR amplicon (Fig. 2B–D, example for *RPL13A*). Due to the very small size of the amplicons (between 71 and 108 bp¹⁶) and to limit the baseline noise always present at the beginning and end of Sanger electropherograms, the Sanger sequencing was performed for both forward and reverse strands (Fig. 2C and D, respectively). This allowed for the successful sequencing of forward and reverse primers and respective probe positions in all HepatoPredict amplicons.

The inclusivity of the RT-qPCR reactions included in the HepatoPredict kit was demonstrated as described in CLSI-MM17 guideline [22]. An inclusivity of 100 % was demonstrated for all targets, except for *DPT* which had an inclusivity of 95 % ([Supplementary File 3](#)).

3.3. Limit of Detection (LoD), limit of blank (LoB), linearity and efficiency

For LoD, LoB, linearity and efficiency determination of each RT-qPCR reaction included in HepatoPredict kit, an RNA pool of FFPE HCC samples was used in accordance with MM16-A guideline [23]. The LoD, for each of the 8 targets included in the HepatoPredict kit, was determined based on ISO 20395:2019, CLSI-MM17, MIQE, and CLSI-EP17-A guidelines [20–22,24]. The Probit model was used, and Fig. 3 exemplifies the application of the model to the *RPL13A* target. The Probit model was applied to two different data sets obtained using different reagent lots (Figs. 3A and 4B). The LoD was defined as the lowest concentration of target that could be detected in ≥95 % of the samples, as represented in Figs. 3A and 4B. To determine the Cq value corresponding to the LoD, a linear regression was performed (Fig. 3C). For all targets, Pearson's correlation coefficient (R^2) was higher than 0.90 (data not shown), except for *Chr3* ($R^2 = 0.73$, data not shown). The highest LoD (nucleic acid concentration) between lots was taken as the reported value for the measurement. The LoD of the 7 RNA targets included in the HepatoPredict kit ranged from 34.75 Cq to 36.89 Cq (Table 1). Regarding *Chr3* the LoD was defined at 33.95 Cq (Table 1).

To apply the Probit model, it is necessary to assume that all blank or negative samples are reported as negative. These assumptions are true for the HepatoPredict kit, in fact, if a valid Cq value was obtained (<40 cycles) for just one of the replicates of the NTC (no template control) the entire assay was considered invalid. Thus, the LoB was assumed to be zero, i.e., Cq results for the NTC samples,

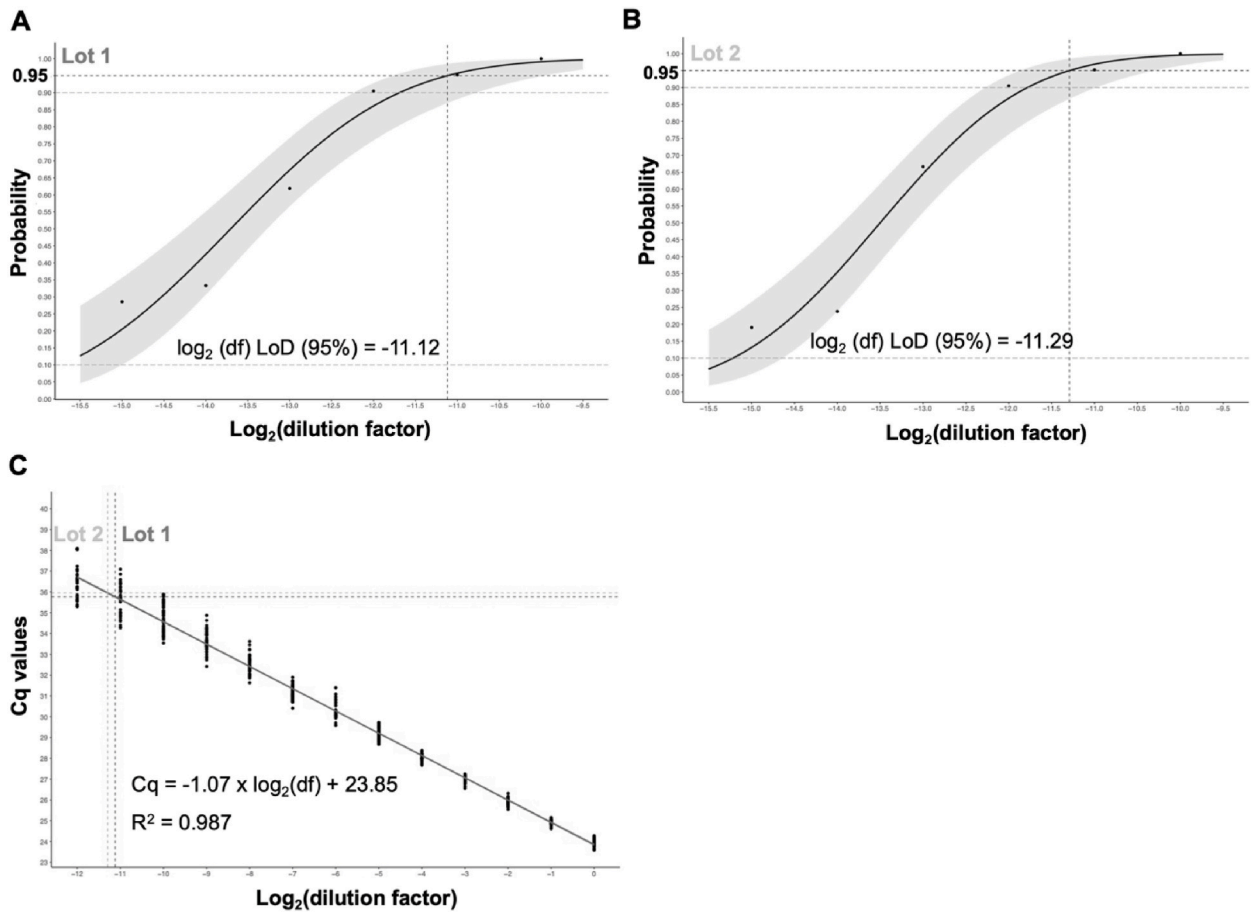


Fig. 3. Estimation of the Limit of Detection (LoD) in Cq values (for *RPL13A* as an illustrative example). The Probit approach was used to determine the LoD of *RPL13A* gene for two different lots of reagents. The LoD was defined as the concentration (log₂ dilution) at a probability of 95 %. The grey areas represent the confidence interval at 99 % (A and B). The linear dynamic range was also estimated and the Cq value associated with the LoD was determined (C). For the example of the *RPL13A* gene, the highest LoD was obtained with Lot 1 and the correspondent Cq value was 35.75.

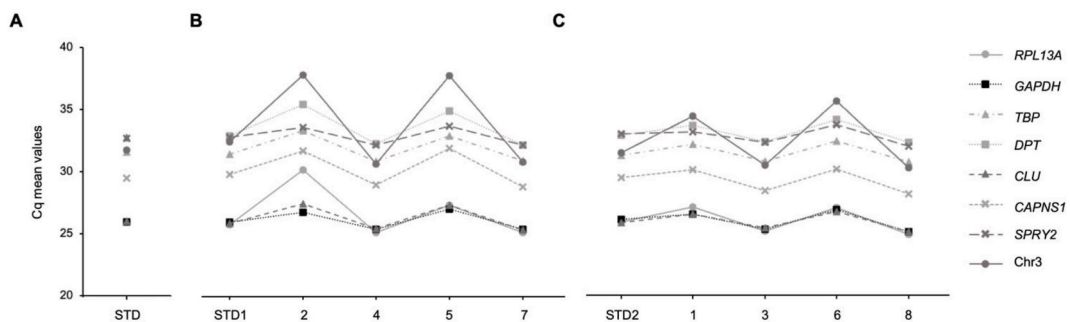


Fig. 4. Robustness of the RT-qPCR reactions included in the HepatoPredict kit. Representation of Cq mean values for each condition (Supplementary File 1) of each target included in the HepatoPredict kit. The original condition (STD) (A) was compared with all other conditions (1–8) and no statistically significant differences were observed (Dunn’s multiple comparisons test). Moreover, conditions with the same annealing temperature were compared with the respective STD condition: (B) STD1 was compared with condition 2, 4, 5 and 7 while (C) STD2 was compared with conditions 1, 3, 6, and 8. No statistically significant differences were observed (Dunn’s multiple comparisons test).

for all valid RT-qPCR reactions, did not cross the threshold within the 40 cycles and were considered “undetermined” (>40 Cq).

The linear range of each target was also determined in accordance with ISO 20395:2019 [20] with R² = 0.99 for *RPL13A*, *GAPDH*, *DPT*, and *CAPNS1*, R² = 0.98 for *TBP* and *CLU*, R² = 0.97 for *SPRY2* and R² = 0.93 for *Chr3*. All targets were linear at least within 24.63

Table 1

LoB, LoD, linear range and PCR efficiency of the 8 targets included in HepatoPredict kit.

Target	LoB Cq value	LoD Cq value	Linear Range		Efficiency %
			Cq value	R [2]	
RPL13A	>40	35.75	17.63–36.30	0.99	91.02
GAPDH	>40	36.32	18.02–38.70	0.99	104.76
TBP	>40	34.75	24.63–39.62	0.98	110.01
DPT	>40	35.11	24.41–37.68	0.99	102.97
CLU	>40	35.35	19.07–39.44	0.98	110.26
CAPNS1	>40	36.89	21.22–38.44	0.99	101.79
SPRY2	>40	34.93	23.27–38.02	0.97	97.70
Chr3	>40	33.95	23.90–35.11	0.93	91.02

and 35.11 Cq (Table 1). Moreover, the amplification efficiencies, determined in accordance with ISO 20395:2019 and MIQE guidelines [20,21], ranged from 91.02 to 110.26 Cq for all targets (Table 1).

Considering that the HepatoPredict kit analyzes 7 different genes and a DNA-specific target, it is important to assure that all targets can be detected within their linear range and below their LoDs. As represented in Table 2, the maximum RNA input to assure an ideal performance of the HepatoPredict kit is 0.031 ng/μL, corresponding to Cq values of 28.26 for RPL13A, 28.34 for GAPDH, 33.95 for TBP, 34.44 for DPT, 29.42 for CLU, 31.61 for CAPNS1, and 33.96 for SPRY2.

Robustness of RT-qPCR reactions: To study the robustness of the RT-qPCR reactions included in the HepatoPredict kit, alterations in the concentrations and volumes of RT-qPCR reagents were performed as suggested in ISO 20395:2019 [20] and represented in Supplementary File 1. The Cq mean values of two independent assays obtained for each target under each condition are represented in Fig. 4A–C. Conditions with the same annealing temperature were compared with the respective standard condition – conditions 2, 4, 5, and 7 were compared with STD1 (Fig. 4B), while conditions 1, 3, 6, and 8 were compared with STD2 (Fig. 4C) – and no statistically significant differences were observed ($p > 0.05$). Furthermore, all conditions (from 1 to 8) were compared with the STD condition representing no changes regarding the initial protocol (Fig. 4A) and no statistically significant differences were observed ($p > 0.05$).

3.4. Precision of the HepatoPredict kit

The precision of the HepatoPredict kit was determined as described in ISO 20395:2019, MIQE, and CLSI-MM17 guidelines [20–22]. Precision data was transduced numerically using imprecision values such as standard deviation (SD) and respective confidence interval (CI) at 95 % (Table 3). The HepatoPredict kit reproducibility was verified by normalizing the gene expression level of the genes of interest (DPT, CLU, CAPNS2, and SPRY2) to the geometric mean of the reference genes (RPL13A, GAPDH, and TBP) (Table 3), as described for the standard use of the kit. In general, the SD for daily, lot-to-lot, and operator reproducibility were higher for the DPT gene (0.38–1.36) when compared with the other genes of interest included in the HepatoPredict kit (0.03–0.44). The same was verified for the inter-assay reproducibility (SD calculated between all the independent assays for the same sample) and total SD (square root of the daily, lot-to-lot, and operator variances) (Table 3). Furthermore, while all targets of sample A were associated with higher SD in

Table 2

Acceptable maximum Cq values for each HepatoPredict kit's target to be detected below their LoD and within the linear range.

Reference RNA (ng/μL)	Targets (Cq mean)						
	RPL13A	GAPDH	TBP	DPT	CLU	CAPNS1	SPRY2
8	20.19	20.40	25.67	25.95	21.38	23.43	25.15
4	21.14	21.36	26.59	26.94	22.44	24.45	26.30
2	22.28	22.37	27.66	27.85	23.42	25.46	27.31
1	23.33	23.27	28.82	28.89	24.44	26.45	28.26
0,5	24.06	24.28	29.76	30.02	25.40	27.68	29.26
0.25	25.09	25.33	30.67	31.12	26.58	28.70	30.50
0.125	26.12	26.43	31.93	32.04	27.40	29.46	31.34
0.063	27.06	27.47	32.76	32.47	28.60	30.75	32.58
0.031	28.26	28.34	33.95	34.44	29.42	31.61	33.96
0.016	29.12	29.38	35.47	35.64	30.53	32.86	34.33
0.008	30.31	30.34	35.31	36.68	31.79	33.89	37.39
0.004	31.20	31.54	36.40	N/A	32.67	35.02	36.80
0.002	32.36	32.41	N/A	38.06	33.68	37.12	N/A
0.00100	33.22	33.33	37.71	N/A	34.57	37.10	38.47
0.00050	34.41	34.18	N/A	N/A	35.20	38.19	N/A
0.00024	35.65	36.31	N/A	37.19	37.15	39.17	N/A
0.00012	36.85	37.37	N/A	N/A	N/A	38.75	N/A

Cq mean values represented in bold are above the LoD of each target. Between double lines are represented the acceptable maximum Cq values for all targets assuring that all of them are detected below their LoD and within their linear range: [RNA] = 0.031 ng/μL. N/A – not applicable, Cq > 40.

Table 3
Reproducibility of the HepatoPredict kit.

Sample	Target	Repeatability (mean SD)	Daily Reproducibility		Lot-to-Lot Reproducibility		Operator Reproducibility		Inter-Assay Reproducibility		Total SD
			SD	95 % CI	SD	95 % CI	SD	95 % CI	SD	95 % CI	
Sample A	DPT	0.07	0.52	0.10–0.93	1.05	0.21–1.89	0.81	0.16–1.46	0.77	0.55–0.99	1.42
	CLU	0.07	0.11	0.02–0.19	0.29	0.06–0.53	0.04	0.01–0.07	0.28	0.20–0.35	0.31
	CAPNS1	0.07	0.24	0.05–0.43	0.35	0.07–0.64	0.17	0.04–0.37	0.26	0.19–0.34	0.46
	SPRY2	0.13	0.08	0.02–0.14	0.08	0.02–0.14	0.11	0.02–0.19	0.18	0.13–0.23	0.16
Sample B	DPT	0.12	0.38	0.08–0.68	0.58	0.12–1.05	0.58	0.12–1.04	0.81	0.58–1.04	0.90
	CLU	0.08	0.34	0.07–0.60	0.30	0.06–0.53	0.19	0.04–0.35	0.25	0.18–0.32	0.49
	CAPNS1	0.11	0.08	0.02–0.14	0.04	0.01–0.07	0.15	0.03–0.26	0.13	0.09–0.16	0.17
	SPRY2	0.08	0.03	0.01–0.06	0.12	0.02–0.21	0.29	0.06–0.52	0.23	0.17–0.30	0.32
Sample C	DPT	0.05	1.36	0.27–2.45	0.41	0.08–0.74	0.84	0.17–1.51	0.99	0.71–1.26	1.65
	CLU	0.14	0.31	0.06–0.55	0.27	0.05–0.48	0.05	0.01–0.09	0.36	0.26–0.47	0.41
	CAPNS1	0.09	0.10	0.02–0.17	0.04	0.01–0.07	0.02	0.00–0.03	0.12	0.09–0.15	0.11
	SPRY2	0.08	0.44	0.09–0.79	0.16	0.03–0.28	0.38	0.08–0.69	0.47	0.34–0.60	0.60
All samples (n=18)	DPT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.69	0.53–0.84	N/A
	CLU								0.26	0.20–0.31	
	CAPNS1								0.24	0.18–0.29	
	SPRY2								0.20	0.16–0.25	

SD, standard deviation; CI, confidence interval; N/A, not applied.

lot-to-lot reproducibility, sample C presented higher SD values in daily reproducibility. Nevertheless, none of these were observed in sample B, suggesting that the observed variability between independent assays is not dependent on a single factor. In fact, when all varying factors were considered (inter-assay reproducibility and total SD), SD values were similar between both samples. Additionally, inter-assay reproducibility was also determined for 18 different HCC FFPE samples (calculation of the mean SD and respective confidence interval), corroborating the previous results (Table 3). The repeatability was verified for each target included in the HepatoPredict kit – reference genes (Supplementary File 5) and genes of interest (Table 3) – and similar SD were obtained for all targets ranging from 0.05 to 0.14 Cq.

3.5. Univariate and multivariate analysis of the new version of the HepatoPredict algorithm

The HepatoPredict algorithm (V2.0) is a two-level predictor, but the first level (Class I), presents the highest precision and uses the XGboost model (instead of SVM in V1.0) increasing the positive predictive value (PPV) to 96.43 %. The second level (Class II) is a linear SVM model. The variables' weights within each model are represented in Supplementary File 6, with DPT and CLU gene expression

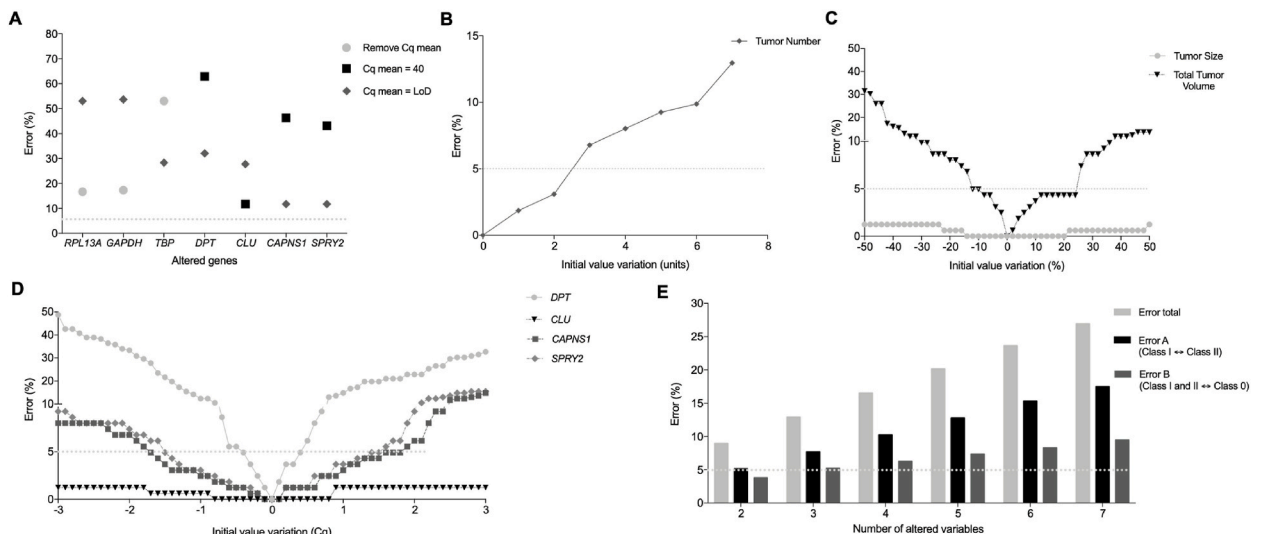


Fig. 5. – Robustness of the HepatoPredict algorithm. Representation of errors associated with the univariate alteration of Cq mean values from each molecular variables before normalization (A), with the univariate alteration of the tumor number (B), or tumor size or total tumor volume (C). Errors associated with the univariate alteration of the molecular variables after normalization (D). Multivariate analyses of the HepatoPredict algorithm for distinct combinations of variables, ranging from 2 to 7 variables (E). The dashed line in light grey represents the errors at 5 %.

levels and total tumor volume being the most important variables of the HepatoPredict algorithm. Furthermore, the new HepatoPredict algorithm was also compared with other clinical criteria for the identification of HCC patients suitable for liver transplantation (Supplementary File 7).

The analytical validation of the new HepatoPredict algorithm consisted in calculating the error (i.e., how many times the correct classification is missed) associated with the alteration of one (univariate) or more (multivariate) variables. Fig. 5A represents the error associated with the alteration of the Cq mean values before normalization. The errors were all superior to 10 % demonstrating that to maintain a good HepatoPredict performance, no reference gene can be removed from the assay and the Cq means of the genes cannot be replaced by 40 Cq or its respective LoD. Moreover, the variations allowed for each variable, assuring a maximum error of 5 % (or a maximum variation of 2 Cq or 50 % from the initial value), are represented in Fig. 5B–D. In general, it was verified that the *DPT* normalized gene expression level and the total tumor volume (in cm³) were the variables that least allowed for alterations (for an error = 5 %, variation of ± 0.4 Cq and ± 12 %, respectively). On the other hand, *CLU* normalized gene expression level and the tumor size (the diameter of the largest tumor in cm) were the variables that tolerated greater variations (± 3 Cq and ± 50 % respectively, while maintaining an error <5 %) (Fig. 5C–D). For these variables, the error at 5 % was not used as a threshold for the multivariable analysis, but instead an acceptable absolute variation value was used (± 2 Cq for *CLU* expression level and ± 50 % for tumor size). Furthermore, while maintaining an error <5 %, the tumor number was possible to vary in 2 units (Fig. 5B) and the normalized gene expression levels of *CAPNS1* and *SPRY2* varied ± 1.7 Cq and ± 1.5 Cq, respectively (Fig. 5D).

The HepatoPredict algorithm multivariate analysis was based on the results presented in Fig. 5B–D and demonstrated that the error increases with the number of varying variables (Fig. 5E). Moreover, two types of errors were analyzed concerning the final HepatoPredict class. Error A corresponds to class switching between HepatoPredict classes associated with a very high or high predicted benefit of LT, respectively Class I and Class II [16]. Error B relates to a switch on HepatoPredict class with more impact in the LT benefit, i.e., a switch from a class with no benefit to a class with LT benefit (Class 0 to Class I or Class II) or from a class with LT benefit to a class with no benefit (Class I or Class II to Class 0). In general, error A was ~ 1.6 times higher than error B (Fig. 5E).

4. Discussion

The HepatoPredict kit uses an algorithm that combines molecular data (gene expression levels of *DPT*, *CLU*, *CAPNS1*, and *SPRY2*) with clinical variables (tumor number, size of the largest nodule, and total tumor volume) to classify the patients in two different classes associated with the benefit of a liver transplant (Class I – very high confidence, and Class II – high confidence) or in Class 0 (no benefit of liver transplant predicted). Some products already exist in the market focused on the prognostic prediction of different tumors, such as breast [25] and prostate [26], but nothing specific for HCC is available. While the successful clinical validation of HepatoPredict kit using a retrospective cohort was previously published [16], the new HepatoPredict algorithm presents some variation as compared to the original algorithm (Supplementary File 8) which reduces the risk of overfitting with new samples. The new algorithm is associated with a higher PPV in Class I, enabling the identification of individuals with a very high likelihood of good prognosis for LT. This could be particularly beneficial in geographies where the waiting period for LT is very long. In addition, the new HepatoPredict algorithm presents a higher negative predictive value (NPV) when compared with different clinical criteria (such as Milan [4], UCSF [6], Up to seven [11], AFP [7], Metroticket 2.0⁵, TTV [13], and TTV AFP [12] criteria). The higher NPV of HepatoPredict translates into a higher probability of being correct in terms of detecting a bad prognosis patient when the kit result is Class 0 (no benefit of a LT). This reduces the misclassifications of patients that benefit from a LT and can avoid wasting a healthy organ in a patient that very likely will face HCC recurrence. This correct prognosis assignment was also corroborated by the results of multivariate analysis of the HepatoPredict algorithm, demonstrating a higher rate of type A errors (switch between Class I and II with benefit prediction) than type B errors (from good prognosis to bad prognosis and vice-versa). To further corroborate the clinical utility of the HepatoPredict kit additional retrospective studies are being planned enrolling patients' cohorts from different geographic localizations and HCC etiologies and a prospective study (NCT0449983) is currently open and recruiting.

In the context of analytical validation of multi-target genomic assays (such as HepatoPredict kit), no evaluation guidelines covering all the relevant aspects required for the diagnostic setting are available. To fill this gap, different guidelines such as ISO201395-2019²⁰, MIQE [21], CLSI-MM16²³, CLSI-MM17²², and CLSI-EP17-A [24] were followed where applicable, to demonstrate that the HepatoPredict kit is a sensitive, specific, and robust test. Thus, the described analytical validation of the HepatoPredict kit is in accordance with standard assay validation processes. This type of approach has been previously used to validate similar prognostic [27–30] and diagnostic [31–33] tests and as a reference for analytical validation of the test in different molecular pathology laboratories.

In diagnostic settings, FFPE is the most used technique for long-term conservation of clinical samples, since it preserves the proteins and vital structures within the tissue while it aids microscopic diagnostic examination, experimental research, and diagnostic/drug development [34]. FFPE samples were thus implemented for the HepatoPredict kit to simplify its adoption by molecular biology and pathology laboratories. The RNA extraction method from FFPE HCC samples was demonstrated to be repeatable (between duplicates) and reproducible (between operators). Regarding gDNA residual contamination, Cq values above 34 in RT-qPCR reactions targeting *Chr3* were observed in 1.15 and 9.19 % of the samples (for operator 1 and 2, respectively). However, the LoD for *Chr3* was determined at 33.95 Cq, meaning that above this Cq value, *Chr3* detection is likely invalid, suggesting that gDNA contamination during RNA extraction from HCC FFPE samples is very residual and approaching 0 % with HepatoPredict kit.

The RNA extracted from FFPE tissues is normally fragmented [35], thus, FFPE sections were digested with heat application (56 °C) and proteinase K to decrease RNA fragmentation and chemical modifications [36,37]. Moreover, the primers of the HepatoPredict kit were designed for the generation of short amplicons to increase gene detection rate [17,38,39], and gene specific reverse transcription and targeted cDNA amplification (1-step RT-qPCR) were performed to increase the accuracy and sensitivity of the RT-qPCR reactions

[40]. Nevertheless, the HepatoPredict kit includes a sample quality control step comprising 1-step RT-qPCR reactions targeting *RPL13A* and *Chr3*. With this procedure, it is possible to determine if the extracted sample contains enough RNA (shown by the Cq value from *RPL13A*) and if gDNA contamination is present (reported by the Cq value from *Chr3*) [17] before proceeding with the kit protocol. This assessment is important because it assures the reproducibility and veracity of the experiments, avoids extra costs associated with the need of repeating the analysis and the wasting of precious tumor samples [21,41]. Considering that HepatoPredict analyses the expression level of 7 different genes, Cq values for each target, allowing all targets to be detected within their linear ranges and below their LoDs, were determined. Thus, an acceptable Cq range for the sample quality control was defined for the *RPL13A* gene: 18.32 to 28.26 Cq. Regarding *Chr3*, all Cq values above its LoD (33.95 Cq) are acceptable since they represent no gDNA contamination.

Regarding RT-qPCR reactions, primer specificity for each target included in the HepatoPredict kit was demonstrated. Although probes (TaqMan® technology) are also included in the RT-qPCR reactions, they were not analyzed in the context of sequencing because they do not amplify PCR products and because their fluorescence is only released in the context of highly specific annealing to the target sequences in the PCR amplicons. Thus, if the Sanger sequencing proves that each primer pair-related PCR amplicon is specific and no unspecific PCR products are detected, each probe can only anneal to the amplified specific product. To further demonstrate that the RT-qPCR reactions included in the HepatoPredict kit could distinguish between target and non-target sequences, an inclusivity of 100 % was demonstrated for all targets, excluding the *DPT* gene which had an inclusivity of 95 %. This result was expected since the downregulation of the *DPT* gene in HCC has been demonstrated and can be already associated with HCC carcinogenesis and progression [42–44]. In contrast with diagnostic systems [32], the exclusivity of the RT-qPCR reactions was not studied since the HepatoPredict kit analyzes the expression level of genes that are not exclusively expressed on HCC cells. Nevertheless, to be analyzed by HepatoPredict kit, each HCC sample needs to be collected by expert clinicians (surgeons or radiologists) and subsequently analyzed at the microscopical level (H&E-stained tissue slides) by a certified pathologist, assuring the specificity of each HCC biopsy submitted to the HepatoPredict kit test. Furthermore, the sensitivity of the RT-qPCR reactions included in the HepatoPredict kit was determined by defining the LoD for each target, as well as the respective linear range. The amplification efficiency was also calculated for all targets being between 90 and 110 % as recommended by ISO 203095:2019.

The robustness of the qPCR reactions included in HepatoPredict kit was studied. It was demonstrated that qPCR reactions were robust, not being affected by small changes either in annealing temperatures and reagents' concentration and volumes.

Moreover, precision studies, assessing both the repeatability and reproducibility of the qPCR reactions included in the HepatoPredict kit, demonstrated that the variability associated with normalized Cq values for *DPT*, *CLU*, *CAPNS1* and *SPRY2* genes was not dependent on a single factor (day, lot, or operator). Furthermore, the inter-assay SD within a sample is similar to the inter-assay SD between 18 samples, demonstrating the reproducibility of the assay independently of the sample used. The *DPT* gene was associated with a higher SD in all assays, nevertheless no differences were observed in *DPT* repeatability in comparison with the other targets. These results suggest that the lower reproducibility of the *DPT* gene (i.e., higher SD), may be associated with *DPT* lower inclusivity (95 %) due to *DPT* downregulation in HCC [42–44].

The robustness of the new HepatoPredict algorithm was also studied and the acceptable variation range for each variable was determined. It was demonstrated that *DPT* gene expression level and total tumor volume were the most sensitive variables. This was expected since in the XGBoost model (first level), the *DPT* gene expression level and total tumor volume had an information gain of 3.75 and 0.29 (respectively), while the other variables had an information gain of zero. Moreover, the SVM model (second level), also corroborated these results since the variables with higher SVM weights were *DPT* gene expression level, total tumor volume, and *CLU* gene expression level. Total tumor volume is related to the number of tumors and tumor diameter measurements; thus, errors in these variables will influence its value. A recent study described a mean error of 0.81 cm when measuring the tumor size using different magnetic resonance imaging pulse sequences [45]. This was reflected in an HepatoPredict type B error of 4.94 %, demonstrating that the prognostic test handles common measuring errors.

5. Conclusions

After testing the HepatoPredict kit with variations that simulate real-life scenarios, it has been determined that the prognostic information provided by the kit remains consistent. Only extreme and unlikely combinations of multiple variables could potentially affect the prognosis. The validation procedures outlined in this study can serve as a guide for other molecular diagnostic laboratories to ensure accurate results with the HepatoPredict test. Furthermore, the analytical testing showed that the HepatoPredict kit can be integrated into routine diagnostic procedures to identify HCC patients who may benefit from a liver transplant, thereby advancing precision medicine.

Funding

This work was supported by the EU EIC Innovation Council (EIC accelerator grant GA No 946364).

CRedit authorship contribution statement

Maria Gonçalves-Reis: Conceptualization, Formal analysis, Investigation, Writing – review & editing. **Daniela Proença:** Investigation, Writing – review & editing. **Laura P. Frazão:** Formal analysis, Writing – original draft. **João L. Neto:** Formal analysis, Writing – review & editing. **Sílvia Silva:** Resources, Writing – review & editing. **Hugo Pinto-Marques:** Resources, Writing – review & editing. **José B. Pereira-Leal:** Writing – review & editing. **Joana Cardoso:** Conceptualization, Validation, Writing – review & editing.

Declaration of competing interest

The work described here is subject to patent WO 2021/064230 A1; JPL, JC, and HPM declare an ownership interest in the company Ophiomics – Precision Medicine.

MGR, DP, LPF, and JLN are employees at Ophiomics – Precision Medicine.

SS has no competing interests.

Data availability

Data will be made available on request.

Acknowledgements

The authors wish to thank to the European Innovation Council for having partially financed this work with a grant under the EIC Accelerator scheme (Contract N°946364). The authors also thank the patients, *Neuralshift*, and the pathology team from the Curry Cabral Hospital headed by António Figueiredo, with a special mention to Clara Rodrigues. Moreover, the authors particularly acknowledge the Biobank IRBLleida (PT20/00021) integrated in the Spanish National Biobanks Network and Xarxa de Bancs de Tumors de Catalunya sponsored by Pla Director d'Oncologia Catalunya (XBTC), as well as the Biobank ISABIAL integrated in the Spanish National Biobanks Network and in the Valencia Biobanking Network for their collaboration.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.plabm.2024.e00365>.

References

- [1] H. Sung, J. Ferlay, R.L. Siegel, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *Ca - Cancer J. Clin.* 71 (3) (2021) 209–249, <https://doi.org/10.3322/caac.21660>.
- [2] M.F. Silva, M. Sherman, Criteria for liver transplantation for HCC: what should the limits be? *J. Hepatol.* 55 (5) (2011) 1137–1147, <https://doi.org/10.1016/j.jhep.2011.05.012>.
- [3] P. Golabi, S. Fazel, M. Otgonsuren, M. Sayiner, C.T. Locklear, Z.M. Younossi, Mortality assessment of patients with hepatocellular carcinoma according to underlying disease and treatment modalities, *Medicine* 96 (9) (2017) e5904, <https://doi.org/10.1097/MD.0000000000005904>.
- [4] V. Mazzaferro, E. Regalia, R. Doci, et al., Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis, *N. Engl. J. Med.* 334 (11) (1996) 693–702, <https://doi.org/10.1056/NEJM199603143341104>.
- [5] V. Mazzaferro, C. Sposito, J. Zhou, et al., Metroticket 2.0 model for analysis of competing risks of death after liver transplantation for hepatocellular carcinoma, *Gastroenterology* 154 (1) (2018) 128–139, <https://doi.org/10.1053/j.gastro.2017.09.025>.
- [6] F.Y. Yao, L. Ferrell, N.M. Bass, et al., Liver transplantation for hepatocellular carcinoma: expansion of the tumor size limits does not adversely impact survival, *Hepatology* 33 (6) (2001) 1394–1403, <https://doi.org/10.1053/jhep.2001.24563>.
- [7] A. Notarpaolo, R. Layese, P. Magistri, et al., Validation of the AFP model as a predictor of HCC recurrence in patients with viral hepatitis-related cirrhosis who had received a liver transplant for HCC, *J. Hepatol.* 66 (3) (2017) 552–559, <https://doi.org/10.1016/j.jhep.2016.10.038>.
- [8] K.J. Halazun, M. Najjar, R.M. Abdelmessih, et al., Recurrence after liver transplantation for hepatocellular carcinoma, *Ann. Surg.* 265 (3) (2017) 557–564, <https://doi.org/10.1097/SLA.0000000000001966>.
- [9] K. Sasaki, D. Morioka, S. Conci, et al., The tumor burden score: a new “metro-ticket” prognostic tool for colorectal liver metastases based on tumor size and number of tumors, *Ann. Surg.* 267 (1) (2018) 132–141, <https://doi.org/10.1097/SLA.0000000000002064>.
- [10] T. Kaido, K. Ogawa, A. Mori, et al., Usefulness of the Kyoto criteria as expanded selection criteria for liver transplantation for hepatocellular carcinoma, *Surgery* 154 (5) (2013) 1053–1060, <https://doi.org/10.1016/j.surg.2013.04.056>.
- [11] J.Y. Lei, W.T. Wang, L.N. Yan, Up-to-seven criteria for hepatocellular carcinoma liver transplantation: a single center analysis, *World J. Gastroenterol.* 19 (36) (2013) 6077–6083, <https://doi.org/10.3748/wjg.v19.i36.6077>.
- [12] C. Toso, G. Meeberg, R. Hernandez-Alejandro, et al., Total tumor volume and alpha-fetoprotein for selection of transplant candidates with hepatocellular carcinoma: a prospective validation, *Hepatology* 62 (1) (2015) 158–165, <https://doi.org/10.1002/hep.27787>.
- [13] C. Macaron, I.A. Hanouneh, R. Lopez, F. Aucejo, N.N. Zein, Total tumor volume predicts recurrence of hepatocellular carcinoma after liver transplantation in patients beyond Milan or UCSF criteria, *Transplant. Proc.* 42 (2010) 4585–4592, <https://doi.org/10.1016/j.transproceed.2010.10.012>.
- [14] N. Mehta, Liver transplantation criteria for hepatocellular carcinoma, including posttransplant management, *Clin. Liver Dis.* 17 (5) (2021) 332–336, <https://doi.org/10.1002/cld.1054>.
- [15] F. Santopaolo, I. Lenci, M. Milana, T.M. Manzia, L. Baiocchi, Liver transplantation for hepatocellular carcinoma: where do we stand? *World J. Gastroenterol.* 25 (21) (2019) 2591–2602, <https://doi.org/10.3748/wjg.v25.i21.2591>.
- [16] H. Pinto-Marques, J. Cardoso, S. Silva, et al., A gene expression signature to select hepatocellular carcinoma patients for liver transplantation, *Ann. Surg.* (2022), <https://doi.org/10.1097/SLA.0000000000005637>. Published online.
- [17] J. Antonov, D.R. Goldstein, A. Oberli, et al., Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and a proper normalization, *Lab. Invest.* 85 (8) (2005) 1040–1050, <https://doi.org/10.1038/labinvest.3700303>.
- [18] R.L. Plackett, J.P. Burman, Biometrika trust the design of optimum multifactorial experiments, *Biometrika* 33 (4) (1946) 305–325, <https://doi.org/10.2307/2332195>.
- [19] H. Pinto-Marques, J. Cardoso, S. Silva, et al., A gene expression signature to select hepatocellular carcinoma patients for liver transplantation, *Ann. Surg.* 276 (5) (2022) 868–874, <https://doi.org/10.1097/SLA.0000000000005637>.
- [20] International Organisation of Standardization. ISO 20395:2019, *Biotechnology – Requirements for Evaluating the Performance of Quantification Methods for Nucleic Acid Target Sequences – qPCR and dPCR*, 2019. ISO 20395:2019. Published online.
- [21] S.A. Bustin, V. Benes, J.A. Garson, et al., The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments, *Clin. Chem.* 55 (4) (2009) 611–622, <https://doi.org/10.1373/clinchem.2008.112797>.

- [22] Clinical & Laboratory Standards Institute (CLSI), Verification and Validation of Multiplex Nucleic Acid Assays, CLSI document MM17-A, 2008. Published online.
- [23] Clinical & Laboratory Standards Institute (CLSI), Use of External RNA Controls in Gene Expression Assays, CLSI document MM16-A, 2006. Published online.
- [24] Clinical & Laboratory Standards Institute (CLSI), Protocols for Determination of Limits of Detection and Limits of Quantitation, NCCLS document EP17-A, 2004. Published online.
- [25] R. Varnier, C. Sajous, S. de Talhouet, et al., Using breast cancer gene expression signatures in clinical practice: unsolved issues, ongoing trials and future perspectives, *Cancers* 13 (19) (2021) 1–24, <https://doi.org/10.3390/cancers13194840>.
- [26] A. Choudhury, C.M.L. West, Translating prognostic prostate cancer gene signatures into the clinic, *Transl. Cancer Res.* 6 (2017) S405–S408, <https://doi.org/10.21037/tcr.2017.03.13>.
- [27] R. Kronenwett, K. Bohmann, J. Prinzler, et al., Decentral gene expression analysis: analytical validation of the Endopredict genomic multianalyte breast cancer prognosis test, *BMC Cancer* 12 (2012), <https://doi.org/10.1186/1471-2407-12-456>.
- [28] M. Cronin, C. Sangli, M.L. Liu, et al., Analytical validation of the oncotype DX genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer, *Clin. Chem.* 53 (6) (2007) 1084–1091, <https://doi.org/10.1373/clinchem.2006.076497>.
- [29] M. Laible, K. Schlombs, K. Kaiser, et al., Technical validation of an RT-qPCR in vitro diagnostic test system for the determination of breast cancer molecular subtypes by quantification of ERBB2, ESR1, PGR and MKI67 mRNA levels from formalin-fixed paraffin-embedded breast tumor specimens, *BMC Cancer* 16 (1) (2016), <https://doi.org/10.1186/s12885-016-2476-x>.
- [30] O. Gruselle, T. Coche, J. Louahed, Development of a quantitative real-time RT-PCR assay for the detection of MAGE-A3-positive tumors, *J. Mol. Diagn.* 17 (4) (2015) 382–391, <https://doi.org/10.1016/j.jmoldx.2015.03.008>.
- [31] N.T. Potter, P. Hurban, M.N. White, et al., Validation of a real-time PCR-based qualitative assay for the detection of methylated SEPT9 DNA in human plasma, *Clin. Chem.* 60 (9) (2014) 1183–1191, <https://doi.org/10.1373/clinchem.2013.221044>.
- [32] J.T. Brown, I.J. Beldorth, W. Laosinchai-Wolf, et al., Analytical validation of a highly sensitive, multiplexed chronic myeloid leukemia monitoring system targeting BCR-ABL1 RNA, *J. Mol. Diagn.* 21 (4) (2019) 718–733, <https://doi.org/10.1016/j.jmoldx.2019.03.002>.
- [33] C. Gürtler, M. Laible, W. Schwabe, et al., Transferring a quantitative molecular diagnostic test to multiple real-time quantitative PCR platforms, *J. Mol. Diagn.* 20 (4) (2018) 398–414, <https://doi.org/10.1016/j.jmoldx.2018.02.004>.
- [34] C. Seiler, A. Sharpe, J.C. Barrett, E.A. Harrington, E.v. Jones, G.B. Marshall, Nucleic acid extraction from formalin-fixed paraffin-embedded cancer cell line samples: a trade off between quantity and quality? *BMC Clin. Pathol.* 16 (1) (2016) 1–13, <https://doi.org/10.1186/s12907-016-0039-3>.
- [35] S. von Ahlfen, A. Missel, K. Bendrat, M. Schlumpberger, Determinants of RNA quality from FFPE samples, *PLoS One* 2 (12) (2007), <https://doi.org/10.1371/journal.pone.0001261>.
- [36] J.Y. Chung, T. Braunschweig, S.M. Hewitt, Optimization of recovery of RNA from formalin-fixed, paraffin-embedded tissue, *Diagnostic Molecular* (2006), <https://doi.org/10.1097/01.pdm.0000213468.91139.2d>. Published online.
- [37] N. Masuda, T. Ohnishi, S. Kawamoto, M. Monden, K. Okubo, Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples, *Nucleic Acids Res.* 27 (22) (1999) 4436–4443, <https://doi.org/10.1093/nar/27.22.4436>.
- [38] K. Specht, T. Richter, U. Mü, A. Walch, M. Werner, H. Hö, Technical advance quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue, *Am. J. Pathol.* 158 (2) (2001), [https://doi.org/10.1016/S0002-9440\(10\)63985-5](https://doi.org/10.1016/S0002-9440(10)63985-5).
- [39] I. Sánchez-Navarro, A. Gámez-Pozo, M. González-Barón, et al., Comparison of gene expression profiling by reverse transcription quantitative PCR between fresh frozen and formalin-fixed, paraffin-embedded breast cancer tissues, *Biotechniques* 48 (5) (2010) 389–397, <https://doi.org/10.2144/000113388>.
- [40] F. Zeka, K. Vanderheyden, E. de Smet, C.A. Cuvelier, P. Mestdagh, J. Vandensompele, Straightforward and sensitive RT-qPCR based gene expression analysis of FFPE samples, *Sci. Rep.* 6 (2016), <https://doi.org/10.1038/srep21418>.
- [41] J.v. Die, B. Román, RNA quality assessment: a view from plant qPCR studies, *J. Exp. Bot.* 63 (17) (2012) 6069–6077, <https://doi.org/10.1093/jxb/err313>.
- [42] Y. Fu, M.X. Feng, J. Yu, et al., DNA methylation-mediated silencing of matricellular protein dermatopontin promotes hepatocellular carcinoma metastasis by $\alpha 3\beta 1$ integrin-Rho GTPase signaling, *Oncotarget* 5 (16) (2014) 6701–6715, <https://doi.org/10.18632/oncotarget.2239>.
- [43] X. Li, P. Feng, J. Ou, et al., Dermatopontin is expressed in human liver and is downregulated in hepatocellular carcinoma, *Biochemistry (Moscow)* 74 (9) (2009) 979–985, <https://doi.org/10.1134/S0006297909090053>.
- [44] S. Liu, J. Qiu, G. He, et al., Dermatopontin inhibits WNT signaling pathway via CXXC finger protein 4 in hepatocellular carcinoma, *J. Cancer* 11 (21) (2020) 6288–6298, <https://doi.org/10.7150/jca.47157>.
- [45] M. Armbruster, M. Guba, J. Andrassy, et al., Measuring hcc tumor size in mri—the sequence matters, *Diagnostics* 11 (11) (2021), <https://doi.org/10.3390/diagnostics11112002>.