Article

The Application of Dynamic Uncertain Causality Graph Based Diagnosis and Treatment Unification Model in the Intelligent Diagnosis and Treatment of Hepatitis B

Nan Deng and Qin Zhang *

School of Computer Science, Beijing Institute of Technology, Beijing 100081, China; annasdn@bit.edu.cn

* Correspondence: zhangqin@buaa.edu.cn

Abstract: Although hepatitis B is widespread, it is hard to cure. This paper presents a new and more accurate model for the diagnosis and treatment of hepatitis B. Based on previous research, the diagnosis and treatment modes were combined into one. By adding more influencing factors and risk factors, the overall diagnosis and treatment model will be further expanded, and a richer and more detailed overall diagnosis and treatment model will be constructed. Reverse logic gates are used in the model to improve the accuracy of the treatment planning. The new unified model is more accurate in subdividing diagnosis results, and it is more flexible and accurate in providing dynamic treatment plans. The prediction process and the static diagnosis process of the model are symmetric, and the related sub-graph is symmetric in structure. In addition, an algorithm for predicting the response probability of treatment scheme is developed, so as to predict the subsequent treatment effects of the current treatment scheme, such as the probability of drug resistance. The results show that this method is more accurate than other available systems, and it has encouraging diagnostic accuracy and effectiveness, which provides a promising help for doctors in diagnosing hepatitis B.

Keywords: AI; DUCG; intelligent diagnosis and treatment; symmetry; hepatitis B

Citation: Deng, N.; Zhang, Q. The Application of Dynamic Uncertain Causality Graph Based Diagnosis and Treatment Unification Model in the Intelligent Diagnosis and Treatment of Hepatitis B. *Symmetry* **2021**, *13*, 1185. <https://doi.org/10.3390/sym13071185>

Academic Editors: Juan Luis García Guirao and Alice Miller

Received: 13 April 2021

Accepted: 23 June 2021

Published: 30 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

HBV infection is a worldwide endemic disease, but its epidemic intensity varies from region to region. Every year, about 887,000 people worldwide die from HBV infection-related diseases, of which 30% are cirrhosis and 45% are primary liver cancer (HCC) [1]. At the same time, with the rapid popularization and development of modern medical service mode characterized by informationization, networking, individuation, and collaboration in the world, the advantages of computer-aided medical diagnosis systems in complex clinical diagnosis and decision support are gradually revealed by virtue of its powerful ability of quantitative data analysis and calculation. Using computer technology to simulate the process of human judgment and reasoning has become an effective aid for doctors in detection, diagnosis, treatment, and prediction [2–4].

Since AI technology can clearly improve the accuracy of diagnosis, provide early warning of disease-related risks, detect changes in patients' physical data, and effectively improve the efficiency of medical diagnosis in the process of assisting clinical diagnosis, the use of AI technology to assist clinical diagnosis has become one of the current research hotspots. Currently, research related to the use of artificial intelligence to assist in clinical diagnosis involves the following areas.

- (1) Medical knowledge mapping and medical terminology standard construction. The standardized representation and structured organization of medical information is the basis for realizing data-driven relevant artificial intelligence technologies to assist medical diagnosis. The related research involves analyzing and processing massive amounts of medical data and draws cognitive conclusions about patients' physical

- status and disease conditions by reasoning, analyzing, comparing, summarizing, and arguing to quickly extract key information from large amounts of data [5,6].
- (2) Artificial intelligence-based medical image analysis. Medical imaging data, as an important diagnostic basis, can accurately and intuitively reflect the disease phenotype, and, coupled with the breakthrough of deep learning technology in image feature extraction, it has become one of the areas where artificial intelligence is currently most closely integrated with assisted diagnosis.
 - (3) Intelligent diagnosis and treatment system. Intelligent diagnosis and treatment system is the application of artificial intelligence technology to the clinical diagnosis process, assisting doctors to complete a series of processes in clinical diagnosis, such as condition collection, diagnosis, and treatment plan recommendation, which is the most important and core application scenario in the medical field. It enables a series of functions, mainly including clinical knowledge collection, clinical medical analysis, treatment effect analysis, error control, cost control, clinical guidance, clinical warning, and efficacy monitoring. In order to solve the problem of intelligent diagnosis of hepatitis diseases, various feature extraction methods have been proposed and used [7,8], most of which have achieved high correctness rates [9].

Dynamic Uncertain Causality Graph (DUCG) is a probabilistic graphical model developed in recent years [10,11]. It is based on domain knowledge, independent of training data, and has strong interpretation ability and robustness, high diagnosis accuracy and high calculation efficiency. Some initial studies to apply DUCG in medical diagnoses were made in Reference [12]. In our previous work, a DUCG-based intelligent diagnosis and treatment system for hepatitis B was developed to support the diagnosis and treatment of hepatitis B. The system achieved a diagnostic accuracy of 99.9% [13].

In terms of the above research, clinical diagnosis has become a new hot spot in the direction of AI research and development [14,15]. However, these current studies have focused on how to improve the correctness of diagnosis and lacked attention to the subsequent treatment plan [16]. Our previous intelligent diagnosis and treatment of hepatitis B system has made a breakthrough in this area by being able to give the current optimal treatment plan after the corresponding diagnostic result [13]. The correct rate of the system for a given treatment scheme reaches 76–95%, which can be said to be a baseline in the intelligent treatment of hepatitis B.

At present, chronic hepatitis B is difficult to be completely cured. After using the treatment scheme, there is also the possibility of adverse reactions. Even if the hepatitis B treatment scheme is effective in treating chronic hepatitis B, it may still produce adverse reactions. Therefore, early prediction of poor treatment effect can guide patients to change the treatment plan in time, make them receive effective treatment earlier, and reduce the risk of adverse reactions and medical expenses. In fact, almost all diagnostic systems at present do not include the ability to provide treatment plans, let alone predict the follow-up effect of treatment plans.

In this article, a new diagnosis and treatment mode of hepatitis B is based on previous research and development. In the new model, the diagnosis and treatment plan of hepatitis B disease is unified as a single entity. This makes the whole diagnosis and treatment mode further expanded and optimized, making the whole system richer and more detailed. A reverse logic gate variable was added to the proposed DUCG. This variable can accurately represent the reverse logic that can not be given by the previous model, helping to improve the correct diagnosis rate of the system. After refining the whole diagnostic and therapeutic model, the Cubic DUCG is introduced, which enables the system to give the treatment plan dynamically. This feature allows the patient's medication to be dynamically monitored, giving timely warning of the need for salvage therapy and giving more flexible and accurate treatment plans. In addition, there is a possibility of adverse reactions to hepatitis B treatment. After a period of treatment, patients have a certain probability of drug resistance or adverse reactions. In view of the above situation, an algorithm was developed to predict the response effect of a treatment plan. For the current given treatment plan,

the system can predict the probability of adverse reactions or drug resistance reactions to the current treatment plan based on real-time examination data of patients. In this way, a reasonable treatment regimen can be targeted to different individual patients in different states, reducing the probability of adverse reactions in patients.

The remainder of this paper is organized as follows:

Section 2 provides the priori knowledge of DUCG. Section 3 describes how the DUCG model provides a dynamic diagnosis and treatment of hepatitis B disease. Section 4 shows a relevant case for predicting response to a treatment regimen. Section 5 concludes this paper and outlines the future work. And the detailed rule description of DUCG can be found in the Appendix A.

2. Introduction to Existing DUCG Theories and Applications

Dynamic Uncertain Causality Graph is a probabilistic graphical model, which describes uncertain causal relationships among multivariate random variables in a graph structure, and brings great convenience to the study of probabilistic models in high-dimensional spaces.

The DUCG models is usually constructed in a modular way, one of which can build local knowledge into a single DUCG modules, and then compile multiple DUCG modules into a complete DUCG expressing the complete knowledge by a DUCG compiler. The compilation process is accompanied by many verification tasks, such as the removal of repeated variables and relationships, parameter consistency verification, etc., so as to ensure the rigor of the synthesized DUCG. The modular model construction method enables DUCG to transform the huge and complex knowledge base construction into many simple local knowledge constructions, which are then automatically synthesized by computer, thus reducing the complexity of model construction and facilitating knowledge management. In addition, the modular construction method makes the DUCG model have good maintainability, and people can arbitrarily replace, delete, or add DUCG modules, as well as compile and synthesize them to form a new DUCG model.

2.1. The Basic Idea of DUCG

The physical meaning and symbolic representation of the variables in DUCG are shown in Table 1, and these variables are used in the construction of DUCG to graphically represent the expert's knowledge. Some terms and nouns in DUCG are shown in Table 2, and these terms and nouns are used in DUCG inference calculations.

The physical meaning and symbolic representation of the variables in DUCG are displayed in Table 1. These variables are used in the construction of DUCG and represent the knowledge of experts in a graphical manner. Some terms and nouns in Dugg are displayed in Table 2, which are used in DUCG reasoning calculations.

Table 1. Meaning and symbolic representation of the variables in the existing DUCG.

Type	Graphics	Meaning
B_i		Type B variables denote root causes, and type B variables have no parent but at least one child variable. During the inference calculation, the B variable is used as the hypothesis to be solved. In the clinical diagnosis model, the B variable is used to represent the disease.
B_i		The double-sided box rectangle B variable here represents a non-primary event. Non-primary events are such events that can exist before a system state abnormality, but do not function until triggered by a primitive event. The bilateral box rectangle B variable in the treatment model represents the treatment plan.
X_i		X-type variables represent outcomes and can also be used as causes. X variables have at least one parent. For example, in clinical diagnosis, X variables are used to indicate symptoms, signs, test results, etc.
BX_i		Integrated cause variable for representing a collection of multiple root causes.

Table 1. Cont.

Type	Graphics	Meaning
D_i		Type D variables indicate default cause events or unknown cause events. Each variable in DUCG, except for the B variable, can be traced back to its cause event (parent event), and, when the cause event of an event is missing or is not of interest, the type D variable is used as its cause event.
G_i		Logical gate variables, used to represent the logical combination between variables, the logical combination relationship of variables is recorded in the logical gate description table LGS_i .
F_{ij}		The weighted event variable, used in DUCG to represent the causal relationship between two variables (V_i, V_j), with the causal effect propagated from the parent variable V_j to the child variable V_i .
conditional F_{ij}		The conditional weight action event variable, which has the same function as the weight action event, is also applied in DUCG to represent the causal relationship between two variables. However, the relationship needs to meet certain conditions to be established, denoted as Z_{ij} . Only when Z_{ij} event is true is the relationship between the two variables established; otherwise, the relationship between the two variables is not established.
P_{ij}		The link event variable, used in DUCG to represent the causal relationship between two variables (V_i, V_j), the causal effect is propagated from the parent variable V_j to the child variable V_i .
conditional P_{ij}		The conditional link event variable is used in DUCG to represent the causal relationship between two variables. The condition Z_{ij} needs to be satisfied for the relationship to be established.

Table 2. DUCG-related terms and meanings.

Terms	Meaning
E'	Abnormal evidence or signals received during the inference diagnosis, $E' = \prod X_{ij}, j \neq 0$
E''	Abnormal evidence or signals received during the inference diagnosis, $E'' = \prod X_{ij}, j \neq 0$
E	All evidence or signals received in the inferential diagnosis, $E = E'E''$
H_{kj}	The hypothetical event to be sought in the inference process, which explains part or all of the anomalous evidence, usually $H_{kj} = B_{kj}$
h_{kj}	To find the posterior probability of the hypothesis H_{kj} , $h_{kj} = \Pr\{H_{kj} E\}$
S_H	The space of possible outcome hypotheses under E conditions in the inference calculation process consists of possible hypotheses H_{kj} , $H_{kj} \in S_H$
ζ_{kj}	E the probability of occurring in a sub-graph containing only the possible hypothesis H_{kj} , $\zeta_{kj} = \Pr\{E sub - DUCG_{kj}\}$

2.2. Application of DUCG in Medical Field

In recent years, research teams have used DUCG to diagnose disease, and some preliminary results have been achieved. One of these is the diagnosis of jaundice. Jaundice is a phenomenon that the concentration of bilirubin in serum increases due to the disorder of bilirubin metabolism, which leads to the yellowing of the sclera, skin, mucosa, other tissues and body fluids. In cooperation with the First Affiliated Hospital of Zhejiang University, we constructed the knowledge base of jaundice diagnosis by screening 27 common diseases with jaundice as the main complaint in clinical diagnosis. After the completion of the knowledge base, medical records were randomly selected from the HIS system of the First Affiliated Hospital of Zhejiang University. Ten medical records are randomly selected for each disease. If the number of medical records in the system was less than 10, all medical

records are used for detection. In the first case, only the patient's risk factors, symptoms and signs in the medical records were entered for testing, and in the second case, all the information of the patient's laboratory examination and imaging examination was entered into the system for testing. The correct rate of diagnosis in the first case was 84.73%, and the second case was 99.01%.

3. DUCG Model for Hepatitis B Disease

According to various test results of hepatitis B patients, there are several clinical diagnosis methods. It is necessary to comprehensively evaluate the disease progress of the patient in order to determine whether antiviral treatment should be started. This includes the patient's serum HBV DNA, ALT levels, and severity of liver disease, as well as age, family history, and concomitant diseases, among others. Therefore, dynamic assessment is more meaningful than a single test in clinical practice.

In the next section, we will describe how the DUCG model can be applied to the diagnosis and dynamic evaluation of chronic hepatitis B.

3.1. Unified Model of Diagnosis and Treatment: Intelligent Diagnosis and Treatment Model for Hepatitis B

Chronic hepatitis B is a serious problem that endangers human health. Without standard treatment, the probability of developing cirrhosis in patients with chronic hepatitis B is about 2–10% per year, and the 5-year cumulative probability of cirrhosis is 12–25%; if patients with compensated cirrhosis are not treated, 3–5% of patients progress to decompensated cirrhosis or liver failure each year; the 5-year survival rate for patients with decompensated cirrhosis or liver failure is only 14–35%, and 65–86% of patients die or require liver transplantation. Normally, patients with hepatitis B first need to go through the following steps to collect relevant information.

1. The patient needs to have a laboratory examination. Detection of hepatitis B virus infection includes serum or blood detection, as well as detection of virus antigens (proteins produced by the virus) or antibodies produced by the patient. Laboratory test results can distinguish whether the patient have hepatitis B or not. Meanwhile, hepatitis B virus antigens and antibodies detected in the blood can also predict the therapeutic effect of antiviral drugs and immune system regulator interferon. Once chronic hepatitis B virus infection is confirmed, the patient need to do some tests to detect and measure HBV DNA. These tests are used to evaluate a person's infection and monitor treatment.
2. The serum levels of these enzymes were collected in detail. The activity of liver enzymes is an important index of liver function. For example, the levels of alanine aminotransferase (ALT) and aspartate aminotransferase (AST) in serum generally reflect the degree of damage of hepatocytes.
3. The imaging examinations, such as medical ultrasound and computed tomography.
4. The purpose of collecting liver biopsy results is to evaluate liver inflammation, rule out other liver diseases, and monitor treatment.
5. Diagnosis of chronic hepatitis B also includes symptoms and signs, such as fatigue, abdominal distension, and jaundice.
6. Consider some basic information, such as gender, age, drinking history, mental illness history, and previous medications.

In addition, seven kinds of drugs and related basic medicine licensed to treat hepatitis B infection were introduced in the construction of knowledge base. In fact, the information used in our diagnosis methods is basically based on the objective examination results and subjective descriptions of patients. Through this information, the system can diagnose hepatitis B patients in more detail and, at the same time, make a treatment plan according to the patient's personal situation.

Based on serology, virology, biochemistry, imaging, pathology, and other ancillary findings in patients with chronic HBV infection, several clinical diagnoses can be made, and these hepatitis B diseases are listed in Table 3.

Table 3. The chronic hepatitis B virus infection subdivisions.

No.	Disease
B_1	HBeAg-positive chronic hepatitis B
B_2	HBeAg-negative chronic hepatitis B
B_3	Chronic hepatitis B virus carriers
B_4	Inactivity HBsAg carriers
B_5	Occult chronic hepatitis B
B_6	Compensated hepatitis B cirrhosis
B_7	De-compensated hepatitis B cirrhosis

At present, there is no drug that can kill or destroy the hepatitis B virus, so it is difficult to completely cure hepatitis B virus. Therefore, the main goal of antiviral therapy for hepatitis B is to maximize long-term inhibition of HBV replication, reduce hepatocyte inflammation necrosis and liver fibrous tissue proliferation, delay and reduce the occurrence of liver failure, cirrhosis de compensation, hepatocellular carcinoma and other complications, improve patients' quality of life, and prolong their survival time. For some patients with appropriate conditions, we should seek clinical cure. Therefore, the treatment of hepatitis B needs long-term persistence in order to obtain the best effect.

In order to minimize drug abuse and treat patients, different treatment plans are given for different diseases and the condition of the patient, and these treatment plans are listed in Table 4.

Table 4. The treatment plans of chronic hepatitis B virus infection.

No.	Treatment Plan	Medical Abbreviation
B_{11}	Interferon alpha-2a	IFN-3MIU
B_{12}	Interferon alpha-2a	IFN-5MIU
B_{13}	PEGylated interferon alpha-2a	PEG-135
B_{14}	PEGylated interferon alpha-2a	PEG-180
B_{15}	Lamivudine	LAM
B_{16}	Adefovir Dipivoxil	ADV
B_{17}	Entecavir	ETV
B_{18}	Telbivudine	LdT
B_{19}	Tenofovir Disoprox	TDF

A model for diagnosis and treatment is given as shown in Figure 1.

The definitions of each X variable are shown in Table 5.

Table 5. The variable definitions in the sub-DUCG of HBeAg-positive chronic hepatitis B.

Name	Description
X_{66}	The symptoms and medical signs of HBeAg-positive chronic hepatitis B
X_{45}, X_{47}, X_{49}	
X_{50-53}	
X_{12}	
X_{18}	
X_{16}	
X_{41}	
X_{10}	
X_{58-62}	
X_{67}, X_{69}	
X_{1601}	Age; Medication history
X_{1801}	
X_{1802}

Figure 1. A model for diagnosis and treatment.

3.2. Reverse Logic Gate

In the current treatment model, a new logic gate is combined and described in detail here. In some medical diagnosis cases, specific disease/cause may lead to a series of simultaneous observations of the consequence. These simultaneous observations provide new evidence for diagnosing diseases. As shown in Figure 1, inverted logic gate RG_i is used to represent this situation, which is drawn as $\boxed{RG_i}$. Its input can be one or more variables of type $\{B, X, BX, G\}$. Its output has at least two observable X -type variables. In Figure 2, for the compensated hepatitis B cirrhosis ($B_{6,1}$) and de-compensated hepatitis B cirrhosis ($B_{7,1}$), three observations are useful: $X_{42,1}$ represents an APRI score greater than 2, $X_{48,1}$ represents the presence of ascites, and $X_{50,1}$ represents the presence of hepatic encephalopathy. If $X_{50,0}$, $X_{48,0}$, and $X_{42,1}$ appear simultaneously, the high likelihood is that $B_{6,1}$ is true. If $X_{50,1}$, $X_{48,1}$, and $X_{42,1}$ appear simultaneously, $B_{7,1}$ is more likely.

Figure 2. An illustrative example of the DUCG with reversal logic gate.

Compared with the $\{G-, SG-\}$ -type variables, there are four differences:

1. The logic gate specification LGS_i of RG_i is composed of its output variables instead of its input variables, and its state is determined by the state combination of its output variables. For example, the logic gate specification LGS_4 of RG_1 in Figure 2 is as shown in Table 6. If E shows that X_{50} , X_{48} , and X_{42} are state-known, the state of RG_1 is determined. If some of them are state-unknown, RG_1 is state-unknown, so no new E is observed.

Table 6. The chronic hepatitis B virus infection subdivisions.

State of RG_1	State Expression
$RG_{1,0}$	Remnant
$RG_{1,1}$	$X_{50,0} X_{48,0} X_{42,1}$
$RG_{1,2}$	$X_{50,1} X_{48,0} X_{42,1}$
$RG_{1,3}$	$X_{50,0} X_{48,1} X_{42,1}$
$RG_{1,4}$	$X_{50,1} X_{48,1} X_{42,1}$

2. RG_i is connected to its output variable X_n through directional arc $F_{n,i}$. The encoded parameter $a_{nk,ij}$ in $F_{nk,ij}$ can only be 0 or 1. The assignment rule for $a_{nk,ij}$ is as follows: Rule 21: If X_{nk} exists in the state expression of RG_{ij} , then $a_{nk,ij} = 1$; otherwise, $a_{nk,ij} = 0$. The meaning of Rule 21 can be regarded as that when RG_{ij} is true, the output variable state in the event expression of RG_{ij} is true, i.e., the probability that RG_{ij} causing X_{nk} is 1. According to Rule 21 and Table 6, we have $a_{50,1}$ and $a_{48,1}$ encoded in Figure 2 as shown in Equation (1). They are coded with $F_{50,1}$ and $F_{48,1}$, respectively. Actually, according to Rule 21, such $a_{nk,ij}$ can be automatically determined by LGS_i .

$$a_{50,1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, a_{48,1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (1)$$

It should be pointed out that when RG_i is a parent, $r_{n,i}$ is not required. This will be explained later. Actually, the output a, r-type parameters are not used because they will be deleted in the simplification as shown below.

3. If E confirms RG_{ij} according to LGS_i , RG_{ij} , as a member of E , jointly confirms that RG_{ij} output X_{nm} of RG_i is removed from E (i.e., RG_i has priority to apply), except that the X_{nm} has other input but is not the parent of RG_i . In this exception, the X_{nm} still remains in E , while $F_{n,i}$ has been eliminated. That is why $r_{n,i}$ and $a_{n,i}$ are really not needed. This situation is shown in Figures 3 and 4, which is simplified from Figure 2 and Table 6. The color nodes are indicated in E . In particular, the green node indicates

state-normal variable that is usually a negative evidence.

Figure 3. The simplified DUCG from Figure 2 and Table 6, given $E = X_{10,1}X_{42,1}$ (X_{11} is state-unknown). The new $E = X_{10,1}RG_{1,1}$. The blue color indicates state 2.

Figure 4. The simplified DUCG from Figure 2 and Table 6, given $E = X_{10,1}X_{42,0}$ (X_{11} is state-unknown). The new $E = X_{10,1}RG_{1,0}$.

In Figure 3, as a new member of evidence E , $RG_{1,1}$ is determined according to Table 6 and the observed $X_{10,1}X_{42,1}$, and the new $E = X_{10,1}X_{42,1}RG_{1,1}$, in which $X_{42,1}$ is removed because it has been included in $RG_{1,1}$ and should not be calculated repeatedly. $X_{10,1}$ remains in the new E because it has B_1 as its input that is not the parent of RG_1 . Note that $F_{10,1}$ is still removed. Moreover, X_{11} , $F_{11,6}$ and $F_{11,7}$ are eliminated by applying Rule 7 (Appendix A).

4. If RG_i is unknown, then RG_i is removed, together with its input and output directed arcs $F_{i,j,k}$ and $F_{n,i}$, because the function of RG_i is defined as the special evidence consisting of a set of observed X_{nm} . For the example of Figure 2 and Table 6, in addition to the fact that X_{42} is state-unknown, if $E = X_{45,1}X_{10,1}$, i.e., X_{50} is state-unknown, RG_1 is state-unknown and RG_1 , $F_{16,1}$, $F_{45,1}$, $F_{42,1}$ and $F_{50,1}$ are removed as shown in Figure 5.

Figure 5. The simplified DUCG from Figure 2 and Table 6, given $E = X_{45,1}X_{10,1}$ (X_{42} and X_{50} are state-unknown). RG_4 and its input and output are removed. There is no new E .

5. If $RG_{i,0}$ is determined, $RG_{i,0}$ will be used as negative evidence, as shown in Figure 4, and similar to 3, its outputs will be removed, except those of parent companies with other input instead of RG_i .
6. After removing RG_i with unknown status, if $X_{nm}(m \neq 0)$ has no input, according to Rule 10 (Appendix A), X_{nm} is taken as the isolated evidence, and imaginary D_n is taken as its input.

3.3. Detailed Reasoning Process of the Treatment Plan

In this paper, we perform the reasoning steps for a disease treatment plan through a case of hepatitis B disease, with the aim of explaining and presenting the entire diagnostic and treatment details, and exploring the validity and correctness of the validation theory. This example is based on the one shown in Figure 6.

Figure 6. Illustrating an example of an inference algorithm.

The relevant parameters involved in the inference calculation are as follows.

$$\begin{aligned}
 sa_{1,1} &= \begin{pmatrix} - & - \\ - & 3 \\ - & 2 \end{pmatrix}, sa_{2,2} = \begin{pmatrix} - & - & - & - \\ - & 0.8 & 6.7 & 5.9 \end{pmatrix}, a_{4,1} = \begin{pmatrix} - & - & - \\ - & 0.5 & 0.8 \\ - & 0.2 & 0.1 \end{pmatrix}, \\
 a_{4,2} &= \begin{pmatrix} - & - \\ - & 0.3 \\ - & 0.8 \end{pmatrix}, a_{6,2} = \begin{pmatrix} - & - \\ - & 0.6 \end{pmatrix}, a_{7,2} = \begin{pmatrix} - & - \\ - & 0.4 \end{pmatrix}, a_{8,2} = \begin{pmatrix} - & - \\ - & 0.7 \end{pmatrix}, \\
 LGS_1 &= \begin{pmatrix} 0 & null \\ 1 & X_{9,1} \end{pmatrix}, LGS_2 = \begin{pmatrix} 0 & null \\ 1 & X_{10,1}X_{11,0} \\ 2 & X_{10,1}X_{11,1} \\ 3 & X_{10,0} \end{pmatrix}, LGS_4 = \begin{pmatrix} 0 & - \\ 1 & X_{6,1}X_{5,0} \\ 2 & X_{6,1}X_{5,1} \end{pmatrix}, \\
 b_1 &= (- \quad 0.2 \quad 0.1)^T, b_2 = (- \quad 0.1)^T, b_3 = (- \quad 0.3)^T
 \end{aligned} \tag{2}$$

where $r_{n,i} = 1$ between all variables, LGS_3 does not exist (no risk factor), and the conditional event of $SA_{3,3}$ is $Z_{3,3} = X_{11,1}$.

Suppose, the relevant evidence $E = X_{5,1}X_{6,1}X_{7,1}X_{9,1}X_{10,1}X_{11,0}$ is collected and X_8 state is unknown. Since evidence E shows that the conditional event $Z_{3,3} = X_{11,1}$ does not hold, $SA_{3,3}$ is deleted. Thus, we can obtain the DUCG graph in Figure 7 after simplification according to evidence E .

In Figure 7, the colorless represents the state unknown, green represents the state normal, and other colors represent the state abnormal.

Based on the states of the relevant risk factors in evidence E , we can determine the states of the relevant special logic gates and then calculate bx .

(1) The risk factor $X_{9,1}$ is observed to be true, and, according to LGS_1 , the state of SG_1 is 1, and $sa_{1,m;1,1}$ applies. According to the formula in Equation (3),

$$bx_{ij} = \Pr\{BX_{i,j}\} = sa_{ij;in}b_{ij}, \tag{3}$$

we have

$$bx_{1,1} = sa_{1,1;1,1}b_{1,1} = 3 \times 0.2 = 0.6, \quad (4)$$

$$bx_{1,2} = sa_{1,2;1,1}b_{1,2} = 2 \times 0.1 = 0.2. \quad (5)$$

(2) The risk factors $X_{10,1}$, $X_{11,0}$ are observed to be true, and, according to LGS_2 , the state of SG_2 is 1, and $sa_{2,m;2,1}$ applies. According to the formula in Equation (3), we have

$$bx_{2,1} = sa_{2,1;2,1}b_{2,1} = 0.8 \times 0.1 = 0.08. \quad (6)$$

Since $bx_{i,0} = 1 - \sum_{j \neq 0} bx_{ij}$, $bx_{i,0}$ does not need to be computed.

Figure 7. Example graph of DUCG after simplification according to evidence E .

After determining the state of SG and computing bx , we perform a further simplification of Figure 7.

(1) B_3 , SG_3 , BX_3 , X_8 , $F_{8,2}$, and $F_{8,3}$ are deleted.

(2) The evidence $X_{5,1}X_{6,1}$ is synthesized as $RG_{4,2}$, when $E = RG_{4,2}X_{6,1}X_{7,1}$, so $X_{5,1}$ is deleted.

(3) The risk factors of B_1 and B_2 are calculated to obtain bx_1 and bx_2 . At this point, the B type variables, risk factors, SG_i , and $SA_{i,j}$ are no longer needed and are reduced and deleted.

At this point, Figure 7 is further simplified to obtain Figure 8.

Figure 8. Figure 7 is an example of DUCG after further simplification.

From Figure 8, we can get $S_H = \{H_{1,1}, H_{1,2}, H_{2,1}\} = \{BX_{1,1}, BX_{1,2}, BX_{2,1}\}$.

Subplotting Figure 8, we can get sub-DUCG₁ and sub-DUCG_{2,1}, as shown in Figure 9 and Figure 10, respectively.

Figure 9. Example of sub-DUCG₁ after splitting.

Figure 10. Example of sub-DUCG_{2,1} after splitting the graph.

Here, it can be observed that in Figure 9, two isolated anomalous evidences appear.

In the industrial system application of DUCG, isolated evidence is identified as irrelevant variables, which are considered as meaningless or spurious signals and should be discarded from the inference calculation. However, in the clinical treatment process, isolated evidence only indicates that it is not related to the disease or medication regimen of that sub-graph, and is a class of evidence that cannot be explained by that sub-graph, and they exist in the sub-DUCG_{ij} of each sub-graph after simplification, and participate in the calculation, so the isolated evidence is meaningful. In the DUCG inference process in medical field, the splitting step is based on the monadic hypothesis of medical diagnosis, the simplified graph is further split by B variables, and the simplified DUCG continues to be split into multiple sub-graphs sub-DUCG_{ij} to show the causal relationship between each hypothesis H_{ij} and the evidence E . And, in sub-DUCG_{ij}, variables that are not causally related to the current hypothesis will become isolated evidence and function to reduce the likelihood of selecting that medication regimen. Therefore, the dashed D variable is used as the cause of the isolated evidence (indicating that the cause is unknown) according to the relevant simplification rules of the DUCG.

To reflect the role of isolated evidence in the diagnosis process, we defined a degree of concern for each outcome variable (X variable) in the DUCG construct to indicate the degree of physician concern for this variable in the clinical diagnosis process.

The attention of the evidence variable V_i is denoted by $\varepsilon_i (1 \leq \varepsilon_i \leq 100)$, where the attention of each state is denoted by $\varepsilon_{in} (1 \leq \varepsilon_{in} \leq 100)$, given by domain experts based on medical knowledge or experience. To express the effect of isolated evidence on the inference result, during the expression expansion, the isolated evidence X_{in} is expanded to the cause dashed line D_i , where the a parameter is defined as the reciprocal of the degree of concern $1/\varepsilon_i$, and the larger the value of ε_i is taken, the greater the value of $1/\varepsilon_i$ is smaller, the more significant the reduction effect on ζ_{ij} is.

In Figure 6, the attention ε_n of all outcome variables is given as $\varepsilon_{n,1} = 10 (n \neq 4)$ and $\varepsilon_{4,1} = \varepsilon_{4,2} = 20$.

According to Figure 9, it is calculated that:

$$\begin{aligned}
\zeta_{1,1} &= \Pr\{E | \text{sub-DUCG}_{1,1}\} \\
&= \Pr\{RG_{4,2}X_{6,1}X_{7,1} | \text{sub-DUCG}_{1,1}\} \\
&= \Pr\{F_{4,2;1,1}BX_{1,1}F_{6,1;6D}D_6F_{7,1;7D}D_7\} \\
&= a_{6,1;6D}a_{7,1;7D}f_{4,2;1,1}bx_{1,1} \\
&= \frac{1}{\varepsilon_{6,1}} \frac{1}{\varepsilon_{7,1}} a_{4,2;1,1}bx_{1,1} \\
&= \frac{1}{10} \times \frac{1}{10} \times 0.2 \times 0.6 \\
&= 0.0012
\end{aligned} \tag{7}$$

$$\begin{aligned}
\zeta_{1,2} &= \Pr\{E | \text{sub-DUCG}_{1,2}\} \\
&= \Pr\{RG_{4,2}X_{6,1}X_{7,1} | \text{sub-DUCG}_{1,2}\} \\
&= \Pr\{F_{4,2;1,2}BX_{1,2}F_{6,1;6D}D_6F_{7,1;7D}D_7\} \\
&= a_{6,1;6D}a_{7,1;7D}f_{4,2;1,2}bx_{1,2} \\
&= \frac{1}{\varepsilon_{6,1}} \frac{1}{\varepsilon_{7,1}} a_{4,2;1,2}bx_{1,2} \\
&= \frac{1}{10} \times \frac{1}{10} \times 0.1 \times 0.2 \\
&= 0.0002
\end{aligned} \tag{8}$$

According to Figure 10, it is calculated that:

$$\begin{aligned}
\zeta_{2,1} &= \Pr\{E | \text{sub-DUCG}_{2,1}\} \\
&= \Pr\{RG_{4,2}X_{6,1}X_{7,1} | \text{sub-DUCG}_{2,1}\} \\
&= \Pr\{F_{4,2;2,1}BX_{2,1}F_{6,1;2,1}BX_{2,1}F_{7,1;2,1}BX_{2,1}\} \\
&= \Pr\{(F_{4,2;2,1} * F_{6,1;2,1} * F_{7,1;2,1})BX_{2,1}\} \\
&= (f_{4,2;2,1} * f_{6,1;2,1} * f_{7,1;2,1})bx_{2,1} \\
&= (a_{4,2;2,1}a_{6,1;2,1}a_{7,1;2,1})bx_{2,1} \\
&= 0.8 \times 0.6 \times 0.4 \times 0.08 \\
&= 0.01536
\end{aligned} \tag{9}$$

where the operator “*” means multiplying the elements of two matrices in the corresponding positions.

Finally, we obtain by calculation:

$$h_{1,1}^s \approx \frac{\zeta_{1,1}}{\zeta_{1,1} + \zeta_{1,2} + \zeta_{2,1}} = \frac{0.0012}{0.0012 + 0.0002 + 0.01536} = 0.07159, \tag{10}$$

$$h_{1,2}^s \approx \frac{\zeta_{1,2}}{\zeta_{1,1} + \zeta_{1,2} + \zeta_{2,1}} = \frac{0.0002}{0.0012 + 0.0002 + 0.01536} = 0.01193, \tag{11}$$

$$h_{2,1}^s \approx \frac{\zeta_{2,1}}{\zeta_{1,1} + \zeta_{1,2} + \zeta_{2,1}} = \frac{0.01536}{0.0012 + 0.0002 + 0.01536} = 0.9165. \tag{12}$$

The ordering of h_{kj}^s is as follows:

$$\begin{aligned}
h_{2,1}^s &= 91.65\% \\
h_{1,1}^s &= 7.159\%. \\
h_{1,2}^s &= 1.193\%
\end{aligned} \tag{13}$$

Thus, the best treatment option is to select the drug $B_{2,1}$.

The entire reasoning process of the treatment protocol is illustrated in detail by this arithmetic example, which explains in detail how the relevant medication regimen, risk factors, combined evidence and isolated evidence are calculated in the reasoning process.

3.4. Dynamic Treatment Plan Reasoning

Because of the development of hepatitis B, the patient's condition is always in a dynamic process. In order to truly meet the needs of the system for dynamic, real-time, high-reliability online diagnosis, and strictly and accurately realize the expression of dynamic processing logic relations and the interpretation and prediction of reasoning results, a new set of cube DUCG theory is called [17]. It systematically investigates how to express the formation, evolution and development laws of hepatitis B disease using stereoscopic cause-effect diagrams, explores the complex causal modeling problems, such as the temporal causality of different treatment options and dynamic response results, and gives an accurate and efficient dynamic inference algorithm.

The general idea of Cubic DUCG research is based on the advantages of Dugg, such as the ability of greatly simplifying knowledge based on observation evidence and the self-dependence of causal reasoning. By allowing causal relations to traverse multiple time slices, a three-dimensional causal diagram is dynamically built online, and then solved based on logic and probabilistic reasoning. Therefore, the dynamic causal evolution of diseases can be visualized, and the future development of diseases can be strictly predicted.

The special terms involved are shown in Table 7.

Table 7. Special terms involved in Cubic DUCG.

Terms	Meaning
t_m	Current diagnosis moment
$Slice_DG(t_m)$	Intra-time-slice DUCG after receiving observed evidence at moment t_m
$Cubic_DG(t_m)$	Overall Cubic DUCG at t_m moments
$Cubic_DG(B_i, t_m)$	The Cubic DUCG indexed by the first causal event B_i at time t_m
$Cubic_DG^*(t_m)$	The inference obtained at moment t_m is shown in the Cubic DUCG
$Cubic_DG^*(B_i, t_m)$	The Cubic DUCG for inference indexed by the first cause event B_i at moment t_m
$CE(t_m)$	The set of all evidence collected at each point in time up to time t_m
$CE(B_i, t_m)$	Evidence covered on $Cubic_DG(B_i, t_m)$ at moment t_m
$S_H(t_m)$	Overall hypothesis space (set of candidate diagnostic root causes)

Assume that evidence $E(t_1)$ is received at time t_1 as follows.

Evidence of anomalies:

$$\begin{aligned} E'_1(t_1) &= X_{10,1}(t_1); E'_2(t_1) = X_{18,1}(t_1); \\ E'_3(t_1) &= X_{1801,2}(t_1); E'_4(t_1) = X_{1802,1}(t_1); \end{aligned} \quad (14)$$

Normal Evidence:

$$\begin{aligned} E''_5(t_1) &= X_{66,0}(t_1); E''_6(t_1) = X_{12,0}(t_1); \\ E'_7(t_1) &= X_{54,0}(t_1); E''_8(t_1) = X_{41,0}(t_1); \\ E''_9(t_1) &= X_{16,0}(t_1); E''_{10}(t_1) = X_{1601,0}(t_1); \\ E''_{11}(t_1) &= X_{1602,0}(t_1); E''_{12}(t_1) = X_{72,0}(t_1); \\ E''_{13}(t_1) &= X_{485,0}(t_1); \end{aligned} \quad (15)$$

In addition, X_{11} and X_{14} are both state unknown variables.

After reasoning about the algorithmic process, intra-temporal slice transversal simplification is performed using the simplification rules. The final sub-graphs obtained are $Slice_DG(B_1, t_1)$ and $Slice_DG(B_2, t_1)$, as shown in Figure 11.

Figure 11. After simplification $Slice_DG(B_1, t_1)$.

Since

$$Cubic_DG^*(B_1, t_1) = Cubic_DG(B_1, t_1) = Slice_DG(B_1, t_1), \quad (16)$$

the logical inference is carried out based on Figure 11, with the included evidence

$$CE(B_1, t_1) = E'(t_1) = X_{10,1}(t_1)X_{18,1}(t_1)X_{1801,2}(t_1)X_{1802,1}(t_1). \quad (17)$$

The evidence variables in $CE(B_1, t_1)$ are logically expanded and the hypothesis space is found as

$$\begin{aligned} E'_1(t_1) &= X_{10,1}(t_1) = F_{10,1,1}(t_1; t_1) = A_{10,1,1}(t_1; t_1)B_1 \\ E'_2(t_1) &= X_{18,1}(t_1) = F_{18,1,1}(t_1; t_1) = A_{18,1,1}(t_1; t_1)B_1 \\ E'_3(t_1) &= X_{1801,2}(t_1) = F_{1801,2,18,1}(t_1; t_1)F_{18,1,1}(t_1; t_1)B_1 = A_{1801,2,18,1}(t_1; t_1)A_{18,1,1}(t_1; t_1)B_1 \\ E'_4(t_1) &= X_{1802,1}(t_1) = F_{1802,1,1801,2}(t_1; t_1)F_{1801,2,18,1}(t_1; t_1)B_1 \\ &= A_{1802,1,1801,2}(t_1; t_1)A_{1801,2,18,1}(t_1; t_1)A_{18,1,1}(t_1; t_1)B_1. \end{aligned} \quad (18)$$

The expansion expression of the anomaly evidence $CE'(B_1, t_1)$ is, thus, obtained as

$$\begin{aligned} CE'(B_1, t_1) &= E'_1(t_1)E'_2(t_1)E'_3(t_1)E'_4(t_1) \\ &= (A_{10,1,1}(t_1; t_1) * A_{1802,1,1801,2}(t_1; t_1)A_{1801,2,18,1}(t_1; t_1)A_{18,1,1}(t_1; t_1))B_1. \end{aligned} \quad (19)$$

So,

$$Pr\{(B_1, t_1)\} = Pr\{E'(t_1)\} = 0.00531. \quad (20)$$

The '*' operation represents the logical operation in the case of intersection of multiple causal links; refer to Reference [18] for details. The expansion expression of $CE'(B_1, t_1)$ leads to the assumption that the space $S_{H_1}(t_1) = \{B_{1,1}\}$, where $H_1 \equiv B_1$ and $H_{1,1} \equiv B_{1,1}$.

Since $Cubic_DG(B_1, t_1)$ does not contain normal evidence, $CE(B_1, t_1) = CE'(B_1, t_1)$ and the state probabilities of the hypothetical events $H_{1,1} = B_{1,1}$ on $Cubic_DG(B_1, t_1)$ at moment t_1 with complete evidence are obtained as follows:

$$h_{1,1}^S(t_1) = h_{1,1}^{S'}(t_1) = 0.786521. \quad (21)$$

Assume that evidence $E(t_2)$ is received at time t_2 as follows.
Evidence of anomalies:

$$\begin{aligned} E'_1(t_2) &= X_{10,1}(t_2); E'_2(t_2) = X_{18,1}(t_2); \\ E'_3(t_2) &= X_{1801,1}(t_2); E'_4(t_2) = X_{1802,1}(t_2); \end{aligned} \quad (22)$$

Normal Evidence:

$$\begin{aligned}
 E_5''(t_2) &= X_{66,0}(t_2); E_6''(t_2) = X_{12,0}(t_2); \\
 E_7'(t_2) &= X_{54,0}(t_2); E_8''(t_2) = X_{41,0}(t_2); \\
 E_9''(t_2) &= X_{16,0}(t_2); E_{10}''(t_2) = X_{1601,0}(t_2); \\
 E_{11}''(t_2) &= X_{1602,0}(t_2); E_{12}''(t_2) = X_{72,0}(t_2); \\
 E_{13}''(t_2) &= X_{485,0}(t_2);
 \end{aligned} \tag{23}$$

In addition, X_{11} and X_{14} are both state unknown variables. $Slice_DG(B_1, t_2)$ is shown in Figure 12 after the horizontal simplification of $Origin_DG(B_1)$ based on $E(t_2)$.

Figure 12. After simplification $Slice_DG(B_1, t_2)$.

The $Cubic_DG(B_1, t_2)$ and $Slice_DG(B_1, t_2)$ are connected vertically and reduced to obtain $Cubic_DG(B_1, t_2)$ as in Figure 13. This oscillation is caused by the combination of the anomalous signal $X_{1801,1}$ with the non-primary event $B_{14,1}$ through the logic gate $G_{14,3}$, and the mechanism of this negative feedback phenomenon is presented in Figure 13.

Figure 13. After simplification $Cubic_DG(B_1, t_2)$.

The nodes that are overlapped together include

$$X_{10,1}(t_1) = X_{10,1}(t_2), X_{18,1}(t_1) = X_{18,1}(t_2), X_{1802,1}(t_1) = X_{1802,1}(t_2). \tag{24}$$

The expression for the expansion of the anomaly evidence on $Cubic_DG * (B_1, t_2)$ is

$$\begin{aligned}
 CE'(B_1, t_2) = & \frac{1}{3} [A_{1802,1;1801,1}(t_2; t_2) A_{1801,1;18,1}(t_2; t_2) A_{18,1;1,1}(t_1; t_1) A_{60,0;60,1}(t_2; t_2) \\
 & A_{58,0;58D}(t_2; t_2) A_{1801,1;1801,2}(t_2; t_2) A_{10,1;1,1}(t_1; t_1) + A_{1802,1;1801,1}(t_2; t_2) \\
 & A_{1801,1;18,1}(t_2; t_2) A_{18,1;1,1}(t_1; t_1) A_{58,0;58D}(t_2; t_2) A_{1801,1;1801,2}(t_2; t_2) \\
 & A_{10,1;1,1}(t_1; t_1) + A_{1802,1;1801,1}(t_2; t_2) A_{1801,1;18,1}(t_2; t_2) A_{18,1;1,1}(t_1; t_1) \\
 & A_{60,0;60,1}(t_2; t_2) A_{1801,1;1801,2}(t_2; t_2) A_{10,1;1,1}(t_1; t_1)] B_{1,1} B_{14,1} + \\
 & \frac{1}{3} [A_{1802,1;1801,1}(t_2; t_2) A_{1801,1;18,1}(t_2; t_2) A_{18,1;1,1}(t_1; t_1) A_{60,0;60,1}(t_2; t_2) \\
 & A_{58,0;58D}(t_2; t_2) A_{1801,1;1801,2}(t_2; t_2) A_{10,1;1,2}(t_1; t_1) + A_{1802,1;1801,1}(t_2; t_2) \\
 & A_{1801,1;18,1}(t_2; t_2) A_{18,1;1,1}(t_1; t_1) A_{58,0;58D}(t_2; t_2) A_{1801,1;1801,2}(t_2; t_2) \\
 & A_{10,1;1,2}(t_1; t_1) + A_{1802,1;1801,1}(t_2; t_2) A_{1801,1;18,1}(t_2; t_2) A_{18,1;1,1}(t_1; t_1) \\
 & A_{60,0;60,1}(t_2; t_2) A_{1801,1;1801,2}(t_2; t_2) A_{10,1;1,2}(t_1; t_1)] B_{1,2} B_{14,1}
 \end{aligned} \quad (25)$$

The hypothesis space on $Cubic_DG(B_1, t_2)$ can be obtained from $CE'(B_1, t_2)$ as

$$SH_1(t_2) = \{B_{1,1} B_{14,1}, B_{1,2} B_{14,1}\}, \quad (26)$$

such that

$$\begin{aligned}
 H_1 & \equiv B_1 B_{14} \\
 H_{1,1} & \equiv B_{1,1} B_{14,1}, \\
 H_{1,2} & \equiv B_{1,2} B_{14,1}, \\
 Pr\{CE'(B_1, t_2)\} & = 0.9743.
 \end{aligned} \quad (27)$$

Further,

$$\begin{aligned}
 Pr\{H_{1,1} CE'(B_1, t_2)\} & = 0.0221 \\
 Pr\{H_{1,2} CE'(B_1, t_2)\} & = 0.9134.
 \end{aligned} \quad (28)$$

This example shows the case where the primary causal event $B_{1,k}$ acts in parallel with the non-primary event $B_{14,1}$ to act as the root variable. Therefore, $B_{1,2} B_{14,1}$ was finally diagnosed as the only valid treatment option under $Cubic_DG(t_2)$, which is also the combined diagnostic conclusion from the inference calculations up to the time of t_2 in this example.

4. Reasoning and Prediction of Response Effects of Treatment Regimens

This paper introduces a research on the prediction method of treatment effects in the current system, aiming at quantifying the patient's response after adopting the treatment scheme as a method to evaluate the advantages of the scheme. Since hepatitis B virus is very difficult to clear, the treatment process can be very long. And, during the treatment process, there is a certain probability of viral resistance [19–21] and adverse reactions [22], and even poor response in some patients [23,24]. Drug resistance is one of the major problems in long-term drug treatment of chronic hepatitis B. Drug resistance can lead to virological breakthrough, biochemical breakthrough, virus rebound, and aggravation of hepatitis. A few patients may have liver failure, acute liver failure, and even death. The development of drug resistance often needs to substitute or add other effective drugs for remedial treatment. Therefore, at the beginning of treatment, care should be taken to prevent the development of drug resistance, and appropriate treatment options should be selected. The problem is that it is difficult to determine the final effect of a treatment plan, unless the risk of the plan can be predicted in time during monitoring period, so that doctors can take appropriate and timely measures.

In the proposed study, the uncertainty of patient variation and timely monitoring of current treatment scheme can promote the quantitative evaluation of the therapeutic effects of current treatment scheme. The model of hepatitis B treatment studied in this paper can vividly describe the complicated causal relationship between symptoms and risk factors caused by the hepatitis B virus, as well as provide a real picture of the disease mechanism and its dynamic changes. Therefore, this model can accurately and reliably predict the possible adverse reactions and clinical drug resistance of patients.

Using evidence-based causal logic reasoning and a rigorous probabilistic calculation method, the model for predicting the risk of drug resistance is constructed as shown in Figure 14. A simple example of a 1-step predictive cause-effect diagram is given here.

Figure 14. An example of a 1-step predictive cause-effect graph.

Assume that $a_{1601,2;16,1} \neq 0$, $a_{1601,0;16,1} = a_{1601,1;16,1} = 0$, and evidence $X_{16,1}$ is considered, so that $X_{1601,2}$ can be uniquely determined in the next step of the disease prediction by following the causal chain. Now, according to the diagnosis and N-step prediction chart (N is an adjustable parameter), under the assumptions of pre-existing symptoms and treatment plans, the risk of LAM drug resistance is estimated. Figure 14 can be considered as a simple example of a mixed diagnosis and prediction causality plot for a case, with the variable X_r representing the LAM resistance event by emergence.

The concept of generalized evidence (expressed by the symbol Ξ) is introduced here to encompass all observed and predicted variables in the diagnostic and predictive graphs, such as $X_{18,1}$, $X_{16,1}$, $X_{1601,2}$, etc., in this case. X_{g,η_g} stands for all variables that have an arc directly connected to X_r , such as $X_{18,1}$ and $X_{1601,2}$.

It is assumed here that (1) all predictor variables have deterministic states, and (2) $\langle X_{g,\eta_g} \rightarrow X_{r,1} \rangle$ exists, and all arcs of $F_{r,1;g,\eta_g} = 0$ have been simplified away. This assumption is reasonable and can be satisfied by iterative causal simplification.

It is now desired to calculate the risk of a drug resistance event (i.e., as $R(X_{r,1})$), in this case, under the current disease hypothesis (e.g., $H_{k,j} = B_{15,1}$ in this case) and concomitant evidence (i.e., Ξ), as in Equation (29).

$$R(X_{r,1}) = \Pr\{X_{r,1}|H_{k,j}\Xi\} = \frac{\Pr\{X_{r,1}H_{k,j}\Xi\}}{\Pr\{H_{k,j}\Xi\}}. \quad (29)$$

Since

$$\Pr\{X_{r,1}H_{k,j}\Xi\} = \Pr\left\{\sum_g F_{r,1;g,\eta_g}\right\} \Pr\{H_{k,j}\Xi\}, \quad (30)$$

the above formula can be equivalently transformed into a more concise and efficient calculation as

$$R(X_{r,1}) = \Pr\left\{\sum_g F_{r,1;g,\eta_g}\right\} = \sum_g \left(\frac{r_{r,g}}{r_r}\right) a_{r,1;g,\eta_g}. \quad (31)$$

Thus, the fall risk prediction in Figure 14 can be calculated as

$$R(X_{r,1}) = \left(\frac{r_{r,1}}{r_r}\right) a_{r,1;18,1} + \left(\frac{r_{r,3}}{r_r}\right) a_{r,1;1602,2} + \sum_{g \notin \{1,3\}} \left(\frac{r_{r,g}}{r_r}\right) a_{r,1;g,\eta_g}. \quad (32)$$

By graphically displaying the drug resistance of a specific therapeutic drugs, this algorithm can not only quantitatively identify the potential risk of drug resistance in clinical practice but also help doctors to make more appropriate treatment plans.

5. Empirical Evaluations

In order to evaluate the performance of the duck-based intelligent diagnosis and treatment system for hepatitis B, we randomly selected 120 patients with hepatitis B from the medical record database in Beijing Ditan Hospital. Sixty cases were selected for diagnosis and detection of hepatitis B, and 90 cases were selected for treatment scheme detection. Doctors who did not participate in the development of the system confirmed the correct diagnosis. In the process of clinical diagnosis, being able to diagnosis based on incomplete information is very important for intelligent diagnosis methods. Therefore, some available medical evidence was discarded and used for experiments of incomplete information. Specifically, only part of the medical data is uploaded to the system, and no further test data is available. We divided the data into two groups, which detected cases with complete information and cases with incomplete information, respectively, and compared the results of diagnosis and treatment plan with the correct cases. The results are as shown in Table 8.

In the current research, we have developed a DUCG-based clinical diagnosis and treatment system for hepatitis B disease, which supports the dynamic diagnosis of hepatitis B disease and then gives the corresponding treatment plan for the diagnosis results. The whole system covers 7 subdivision directions of hepatitis B disease, and 9 different treatment plans. In addition, 189 variables and 321 directed causal arcs are involved in this system. After the construction of the knowledge base in the system, it can be seen that the reasoning process and the prediction process of the whole model are symmetrical. Clinical data related to disease, such as symptoms, signs, examination results, medical history, gender, age, and environmental factors, are all included in the causal diagnosis model. In addition, data about drug resistance management was included in the treatment protocol model. For this reason, we targeted a modular modeling scheme, which allows the complex causality graph to be divided into a number of semi-independent pieces from multiple perspectives. As a result, the difficulty of building the whole complex system is greatly reduced. We extended the current system to allow the model to dynamically perform clinical diagnosis and to give targeted new treatment regimens based on previous drug regimens. We compared the results obtained from the current system with the previous work, as shown in Table 9.

Remark 1. *The reason why we select only 10 cases for each disease is because most diseases have less than 10 cases. If we increase the test cases for each disease, most of the increased cases would be for only a few common diseases, which would improperly increase the precision because the common diseases are easy to diagnose. On the other hand, 10 cases have covered most knowledge points in the DUCG and the marginal utility decreases greatly.*

Remark 2. *In the actual diagnosis process, there are cases where some of the test results are missing. For example, due to financial constraints, patients did not apply for examinations that are too expensive. In order to simulate these conditions in the real diagnosis process, part of the information in the case has been hidden randomly. The results obtained are divided into two cases: complete medical information and incomplete medical information. Even in the case of incomplete medical information, the correct rate of the system is still acceptable. This result proves the robustness of the system to incomplete information.*

Remark 3. *After adopting the new normalized model, we tested it again, and the correct rate was significantly improved compared to the previous system. Some of the previous incorrect arithmetic cases were also corrected.*

From Table 9, it can be easily concluded that the current system is obviously more

correct in diagnosis and treatment than the previous system. Then, we use the existing clinical data on hepatitis B to verify the system, and the final results are exciting. Our system is efficient, precise, and robust. Thanks to the chain reasoning method and weighted parameters, the whole diagnosis processing model can obtain high diagnosis accuracy and efficiency under the condition of incomplete information.

Table 8. The correct rate of diagnosis and treatment plan in DUCG-based system.

Disease	Incomplete Medical Information			Complete Medical Information		
	Tested	Correct	Precision	Tested	Correct	Precision
HBeAg-positive chronic hepatitis B	10	9	90%	10	9	90%
HBeAg-negative chronic hepatitis B	10	8	80%	10	9	90%
Chronic hepatitis B virus carriers	10	6	60%	10	8	80%
Inactivity HBsAg carriers	4	2	50%	4	4	100%
Occult chronic hepatitis B	10	8	80%	10	9	90%
Compensated hepatitis B cirrhosis	6	4	66.60%	6	5	83.30%
Decompensated hepatitis B cirrhosis	10	8	80%	10	9	90%
Total	60	45	75.00%	60	53	88.3%
Treatment plan						
IFN-3MIU	10	6	60%	10	9	90%
IFN-5MIU	10	7	70%	10	9	90%
PEG-135	10	6	60%	10	9	90%
PEG-180	10	8	80%	10	10	100%
LAM	10	9	90%	10	10	100%
ADV	10	8	80%	10	10	100%
ETV	10	10	100%	10	10	100%
LdT	10	7	70%	10	9	90%
TDF	10	8	80%	10	10	100%
Total	90	69	76.60%	90	86	95.50%

Table 9. Current DUCG system-based diagnostic and treatment plan correctness versus previous.

Disease	Previous Medical Information			Current Medical Information		
	Tested	Correct	Precision	Tested	Correct	Precision
HBeAg-positive chronic hepatitis B	10	9	90%	10	10	100%
HBeAg-negative chronic hepatitis B	10	9	90%	10	10	100%
Chronic hepatitis B virus carriers	10	8	80%	10	9	90%
Inactivity HBsAg carriers	4	4	100%	4	4	100%
Occult chronic hepatitis B	10	9	90%	10	10	100%
Compensated hepatitis B cirrhosis	6	5	83.30%	6	6	100%
Decompensated hepatitis B cirrhosis	10	9	90%	10	10	100%
Total	60	53	88.3%	60	59	98.3%
Treatment plan						
IFN-3MIU	20	18	90%	20	20	100%
IFN-5MIU	20	18	90%	20	19	95%
PEG-135	20	14	70%	20	18	90%
PEG-180	20	18	90%	20	20	100%
LAM	20	20	100%	20	20	100%
ADV	20	17	85%	20	19	95%
ETV	20	18	90%	20	19	95%
LdT	20	16	80%	20	19	95%
TDF	20	19	95%	20	20	100%
Total	180	158	87.8%	180	174	96.7%

In addition, our system can intuitively express the reasoning process with graphics, which makes the diagnosis conclusion and treatment plans easier to understand and accepted, and it improves the persuasiveness of clinical diagnosis suggestion. All these characteristics indicate that our system has medical relevance in the diagnosis and treatment

of hepatitis B, and it is expected to be used in clinical diagnosis. We will use more abundant clinical data to test our system in future work, as well as further evaluate and improve the system to make it more perfect and abundant.

Author Contributions: Conceptualization, Q.Z. and N.D.; methodology, Q.Z.; software, Q.Z. and N.D.; validation, N.D.; formal analysis, Q.Z. and N.D.; data curation, N.D.; writing—original draft preparation, N.D.; writing—review and editing, Q.Z.; funding acquisition, Q.Z. Both authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61402266, Grant 61273330 and Grant 71671103.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Patient consent was waived due to the hospital only provided historical test data and kept patient identification information strictly confidential.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Rule 1: If E shows that $Z_{n,i}$ is not met, $F_{n,i}$ or $P_{n,i}$ is eliminated from the DUCG. If E shows that $Z_{n,i}$ is met, the conditional $F_{n,i}$ or $P_{n,i}$ becomes the ordinary $F_{n,i}$ or $P_{n,i}$.

Rule 2: If E shows that V_{ij} , $V \in \{B, X\}$, is true while V_{ij} is not a parent event of X_n , $F_{n,i}$ or $P_{n,i}$ is eliminated from the DUCG.

Rule 3: If E shows that X_{nk} is true while X_{nk} cannot be caused by any states of V_{ij} , $V \in \{B, X, G\}$, $F_{n,i}$ or $P_{n,i}$ is eliminated from the DUCG, except that V_i is included in a hypothesis, or is a descendant of an event included in a hypothesis, and the causality chain between them is not blocked by a known event.

Rule 4: If E shows that X_{nk} and V_{ij} , $V \in \{B, X\}$, are true while X_{nk} cannot be caused by V_{ij} , $F_{n,i}$ or $P_{n,i}$ is eliminated from the DUCG.

Rule 5: If the state unknown X_n without input variable or G_n without input variable is encountered, X_n or G_n and its output directed arcs are eliminated from the DUCG.

Rule 6: If G_i without any output is encountered for any reason, G_i is eliminated from the DUCG.

Rule 7: If (1) the state of X_n is unknown, (2) X_n does not have any output, and (3) X_n is not predetermined in concern, X_n and all its input directed arcs are eliminated from the DUCG.

Rule 8: If E shows that X_{nk} and V_{ij} , $V \in \{B, X\}$, are true and X_{nk} appears earlier than V_{ij} , which means that V_{ij} cannot be the cause of X_{nk} , the F or P type variables (they are the members of the causality chain from V_{ij} to X_{nk} and are not related to any other upstream causality chain of X_{nk}) are eliminated from the DUCG.

Rule 9: If there is such a group of variables (named as the independent group) that have no causal connection with those variables related to E , and no variable in this group is predetermined in concern, this independent group of variables can be eliminated from the DUCG.

Rule 10: If E shows X_{nk} is true while X_{nk} does not have any input due to any reason, add a virtual parent event D_n to X_{nk} with $a_{nk;nD} = 1$ and $a_{nk';nD} = 0, k \neq k'$. $r_{n;D}$ can be any value.

References

1. WHO. *Global Hepatitis Report 2019-11-06*; WHO: Geneva, Switzerland, 2017.
2. Esfandiari, N.; Babavalian, M.R.; Moghadam, A.M.E.; Tabar, V.K. Knowledge discovery in medicine: Current issue and future trend. *Expert Syst. Appl.* **2014**, *41*, 4434–4463. [[CrossRef](#)]

3. Isola, R.; Carvalho, R.; Tripathy, A.K. Knowledge discovery in medical systems using differential diagnosis, LAMSTAR & k-NN. *IEEE Trans. Inf. Technol. Biomed. Publ. IEEE Eng. Med. Biol. Soc.* **2012**. [[CrossRef](#)]
4. Barabási, A.L.; Gulbahce, N.; Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **2011**, *12*, 56–68. [[CrossRef](#)] [[PubMed](#)]
5. Chen, Y.; Elenee Argentinis, J.; Weber, G. IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clin. Ther.* **2016**, *38*, 688–701. [[CrossRef](#)] [[PubMed](#)]
6. Dilsizian, S.E.; Siegel, E.L. Artificial Intelligence in Medicine and Cardiac Imaging: Harnessing Big Data and Advanced Computing to Provide Personalized Medical Diagnosis and Treatment. *Curr. Cardiol. Rep.* **2014**, *16*, 441. [[CrossRef](#)] [[PubMed](#)]
7. Polat, K.; Güneş, S. An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. *Digit. Signal Process.* **2007**, *17*, 702–710. [[CrossRef](#)]
8. Dogantekin, E.; Dogantekin, A.; Avci, D. Automatic hepatitis diagnosis system based on Linear Discriminant Analysis and Adaptive Network based on Fuzzy Inference System. *Expert Syst. Appl.* **2009**, *36*, 11282–11286. [[CrossRef](#)]
9. Ster, B.; Dobnikar, A. Neural networks in medical diagnosis: Comparison with other methods. In Proceedings of the International Conference on Engineering Applications of Neural Networks, London, UK, 17–19 June 1996; pp. 427–430.
10. Qin, Z. The Difference between Single-Valued and Multi-Valued Cases in the Compact Representation of CPD in Bayesian Networks. In Proceedings of the Complex Sciences, First International Conference Complex, Shanghai, China, 23–25 February 2009; Revised Papers.
11. Yao, Q.; Qin, Z.; Peng, L.; Ping, Y.; Wang, X. Application of dynamic uncertain causality graph in spacecraft fault diagnosis: Logic cycle. In Proceedings of the International Conference on Computer-Aided Design, Irvine, CA, USA, 13–17 November 2017.
12. Hao, S.R.; Geng, S.C.; Fan, L.X.; Chen, J.J.; Zhang, Q.; Li, L.J. Intelligent diagnosis of jaundice with dynamic uncertain causality graph model. *J. Zhejiang Univ. Sci. B* **2017**, *18*, 393–401. [[CrossRef](#)] [[PubMed](#)]
13. Deng, N.; Zhang, Q. Towards Dynamic Uncertain Causality Graphs for the Intelligent Diagnosis and Treatment of Hepatitis B. *Symmetry* **2020**, *12*, 1690. [[CrossRef](#)]
14. Shankaracharya, M.; Kumari, S.; Vidyarthi, A.S. Development of Java based graphical user interface for Diagnosis of Hepatitis Using I Mixture of Expert. In Proceedings of the International Conference in Microbial, Jena, Germany, 5–8 November 2012.
15. Anto, S.; Chandramathi, S.; Aishwarya, S. An expert system based on LS-SVM and simulated annealing for the diagnosis of diabetes disease. *Int. J. Inf. Commun. Technol.* **2016**, *9*, 88. [[CrossRef](#)]
16. Rouhani, M.; Haghighi, M.M. The Diagnosis of Hepatitis Diseases by Support Vector Machines and Artificial Neural Networks. In Proceedings of the Computer Science & Information Technology-Spring Conference, Singapore, 17–20 April 2009.
17. Dong, C.L.; Zhang, Q. Identification of Pivotal Causes and Spreaders in the Time-Varying Fault Propagation Model to Improve the Decision Making under Abnormal Situation. *Qual. Reliab. Eng. Int.* **2016**, *32*, 99–109. [[CrossRef](#)]
18. Qin, Z.; Dong, C.; Yan, C.; Yang, Z. Dynamic Uncertain Causality Graph for Knowledge Representation and Probabilistic Reasoning: Statistics Base, Matrix, and Application. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 645.
19. Tenney, D.J.; Rose, R.E.; Baldick, C.J.; Pokornowski, K.A.; Eggers, B.J.; Fang, J.; Wichroski, M.J.; Xu, D.; Yang, J.; Wilber, R.B.; et al. Long-term monitoring shows hepatitis B virus resistance to entecavir in nucleoside-naïve patients is rare through 5 years of therapy. *Hepatology* **2010**, *49*, 1503–1514. [[CrossRef](#)] [[PubMed](#)]
20. Liu, Y.; Corsa, A.C.; Buti, M.; Cathcart, A.L.; Flaherty, J.F.; Miller, M.; Kitrinis, K.M.; Marcellin, P.; Gane, E.J. No detectable resistance to tenofovir disoproxil fumarate in HBeAg⁺ and HBeAg[−] patients with chronic hepatitis B after 8 years of treatment. *J. Viral Hepat.* **2016**, *24*, 68–74. [[CrossRef](#)] [[PubMed](#)]
21. Lim, Y.S.; Yoo, B.C.; Byun, K.S.; Kwon, S.Y.; Kim, Y.J.; An, J.; Lee, H.C.; Lee, Y.S. Tenofovir monotherapy versus tenofovir and entecavir combination therapy in adefovir-resistant chronic hepatitis B patients with multiple drug failure: results of a randomised trial. *Gut* **2016**, *65*, 1042–1051. [[CrossRef](#)] [[PubMed](#)]
22. Agarwal, K.; Brunetto, M.; Seto, W.; Lim, Y.; Fung, S.; Marcellin, P.; Ahn, S.; Izumi, N.; Chuang, W.; Bae, H. 96 weeks treatment of tenofovir alafenamide vs. tenofovir disoproxil fumarate for hepatitis B virus infection. *J. Hepatol.* **2018**, *68*, 672. [[CrossRef](#)] [[PubMed](#)]
23. EASL. 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. *J. Hepatol.* **2017**, *67*, 370–398. [[CrossRef](#)] [[PubMed](#)]
24. Terrault, N.A.; Lok, A.S.F.; McMahon, B.J.; Chang, K.M.; Hwang, J.P.; Jonas, M.M.; Brown, R.S., Jr.; Bzowej, N.H.; Wong, J.B. Update on Prevention, Diagnosis, and Treatment of Chronic Hepatitis B: AASLD 2018 Hepatitis B Guidance. *Clin. Liver Dis.* **2018**, *12*, 33–34. [[CrossRef](#)] [[PubMed](#)]