

78140_Auto_Edited-check.docx

Risk factor analysis and clinical decision tree model construction for diabetic retinopathy in Western China

Zhou YY *et al.* Clinical prediction model for diabetic retinopathy

Abstract

5

BACKGROUND

Diabetic retinopathy (DR) is the driving force of blindness in patients with type 2 diabetes mellitus (T2DM). DR has a high prevalence and lacks effective therapeutic strategies, underscoring the need for early prevention and treatment. Yunnan province, located in the southwest plateau of China, has a high prevalence of DR and an underdeveloped economy.

AIM

To build a clinical prediction model that will enable early prevention and treatment of DR.

5

METHODS

In this cross-sectional study, 1654 Han population with T2DM were divided into groups without ($n = 826$) and with DR ($n = 828$) based on fundus photography. The DR group was further subdivided into non-proliferative DR ($n = 403$) and proliferative DR ($n = 425$) groups. A univariate analysis and logistic regression analysis were conducted and a clinical decision tree model was constructed.

RESULTS

Diabetes duration ≥ 10 years, female sex, standing- or supine systolic blood pressure (SBP) ≥ 140 mmHg, and cholesterol ≥ 6.22 mmol/L were risk factors for DR in logistic regression analysis (odds ratio = 2.118, 1.520, 1.417, 1.881, and 1.591, respectively). A greater severity of chronic kidney disease (CKD) or hemoglobin A1c increased the risk

of DR in patients with T2DM. In the decision tree model, diabetes duration was the primary risk factor affecting the occurrence of DR in patients with T2DM, followed by CKD stage, supine SBP, standing SBP, and body mass index (BMI). DR classification outcomes were obtained by evaluating standing SBP or BMI according to the CKD stage for diabetes duration < 10 years and by evaluating CKD stage according to the supine SBP for diabetes duration ≥ 10 years.

CONCLUSION

Based on the simple and intuitive decision tree model constructed in this study, DR classification outcomes were easily obtained by evaluating diabetes duration, CKD stage, supine or standing SBP, and BMI.

Key Words: Diabetic retinopathy; Type 2 diabetes; Western China; Decision tree

Zhou YY, Zhou TC, Chen N, Zhou GZ, Zhou HJ, Li XD, Wang JR, Bai CF, Long R, Xiong YX, Yang Y. Risk factor analysis and clinical decision tree model construction for diabetic retinopathy in Western China. *World J Diabetes* 2022; In press

Core Tip: Due to the underdeveloped economy and higher prevalence of diabetic retinopathy (DR), Yunnan province is facing a serious task of prevention. Based on a large sample of the Han population with type 2 diabetes mellitus in Yunnan province, this study constructed a cost-effective predictive model that may facilitate the timely and individualized estimation of DR risk.

INTRODUCTION

Type 2 diabetes mellitus (T2DM), a common chronic disease, poses a severe threat to human health and quality of life. By 2045, an estimated 552 million people worldwide will suffer from T2DM^[1]. Given the socio-economic developments over the past 40 years, China has become increasingly urbanized, with a growing aging population.

Moreover, obesity and overweight caused by lifestyle changes contribute to the growing burden of T2DM in China^[2]. As a major vascular complication of T2DM, diabetic retinopathy (DR) is currently largely responsible for blindness in the working-class^[3,4].

The two major challenges associated with DR include the high disease prevalence and lack of effective treatments. The global prevalence of DR is 34.6%. As of 2011, 126.6 million people suffered from DR, and this number is estimated to reach 191.0 million in 2030 without effective and timely measures^[5]. The concerning prevalence and severity of DR worldwide may be modulated by racial/ethnic disparities, socio-economic status, health care systems, lifestyles, research methods, and other factors^[6]. Regarding Asian populations, the prevalence of DR varies according to the region. For example, the prevalence of DR is 20.1%^[7], 25.7%^[8], and 35.0%^[9] in Chinese Singaporeans, Chinese Americans, and Taiwanese Chinese, respectively. Further, the prevalence of DR in inland areas of China is 23%. It is also higher in rural areas than in urban areas and in northern areas than in southern and eastern coastal areas^[10-12].

Nevertheless, there is currently a paucity of effective treatments for DR. In addition to systematic interventions for controlling blood glucose levels, blood pressure, and blood lipid levels, several modern therapies have been developed, such as laser photocoagulation^[13] and intravitreal injections of anti-vascular endothelial growth factor (VEGF) antibodies or glucocorticoids, which can delay the progress of proliferative DR (PDR)^[14,15]. However, several side effects associated with these therapies should be noted. For instance, photocoagulation may cause potential retinal damage, anti-VEGF injections are associated with relapse after drug withdrawal, and glucocorticoid use contributes to cataracts and elevated intraocular pressure in a considerable number of patients. Moreover, intraocular injections may cause complications such as endophthalmitis, intraocular hemorrhage, vitreous hemorrhage, and even retinal detachment^[14,15]. Therefore, the utilization of these therapeutic options in clinical practice should be based on systematic evaluation and strict indications.

The high prevalence of DR and lack of efficient therapeutic strategies are associated with reduced quality of life of patients and pose a substantial socio-economic burden on individuals, families, and the society^[16]. According to an analysis of the pedigree of T2DM in Yunnan province, the prevalence of DR^[17] approximates the national average^[10]. Therefore, the active search for associated risk factors is a fundamental priority for the prevention of DR. In this regard, a decision tree model established using identified risk factors is a useful tool. Distinct from traditional statistical methods such as logistic regression analysis, decision trees are effective machine-learning algorithms that solve classification problems. This method obtains a set of effective classification rules through systematic learning of multiple attributes of samples with known classification results. When faced with new unknown samples, the choice of classification or characteristic attributes can be quickly obtained based on the set of rules extracted from the established decision tree^[18-20]. Thus, a decision tree is a prediction model with a simple and intuitive flowchart structure that is particularly suitable for use in clinical practice. This study examined the risk factors associated with DR in the Han population with T2DM in Yunnan province and constructed a clinical decision tree model.

MATERIALS AND METHODS

Study subjects

Patients from the Han population with T2DM were enrolled from the Department of Endocrinology, Affiliated Hospital of Yunnan University. All patients fulfilled the Chinese Diabetes Association criteria for the diagnosis of T2DM^[21]. The criteria for exclusion were as follows: (1) Age < 18 years; (2) positive islet autoantibodies [including islet cell autoantibodies and autoantibodies to glutamic acid decarboxylase-65, insulin, the tyrosine phosphatases islet antigen 2 (IA-2) and IA-2 β , and zinc transporter 8]; (3) acute complications of diabetes mellitus (including diabetic ketoacidosis or diabetic hyperosmolar state); (4) severe hepatic damage; (5) malignant tumors; (6) acute or chronic infectious diseases; (7) other eye diseases (including glaucoma, retinal vascular

occlusion, and ischemic optic neuropathy); and (8) pregnancy. Finally, 1654 patients with T2DM were enrolled.

Ethical principles

This study was approved by the Ethics Committee of Affiliated Hospital of Yunnan University (No. 2021062), and written informed consent was obtained from all participants according to the principles of the Helsinki Declaration.

This trial registration was registered at ChiCTR (ChiCTR2100041888; registration on January 9, 2021, <http://www.chictr.org.cn/index.aspx>).

Clinical information collection

Patient sex, age, diabetes duration, height, weight, waist circumference, hip circumference, waist-hip ratio (WHR), body mass index (BMI), and systolic blood pressure (SBP) and diastolic blood pressure (DBP) (both in the standing and supine positions) were recorded.

Laboratory assessments

All patients fasted at 22:00 the day before blood collection. At 8:00 the next day, 6 mL of venous blood was collected. Fasting blood glucose (Glu0), hemoglobin A 1c (HbA1c), serum creatinine (Scr), blood urea nitrogen (BUN), uric acid (UA), triglycerides (Trig), cholesterol (Chol), high-density lipoprotein Chol (HDL-C), and low-density lipoprotein Chol (LDL-C) were measured. Based on the formula $eGFR (mL/min/1.73 m^2) = 175 \times Scr (mg/dL)^{-1.234} \times age^{-0.179} \times (0.790 \text{ for women})^{[22]}$, the estimated glomerular filtration rate (eGFR) was calculated.

Ophthalmological measurements

All patients underwent non-mydriatic fundus photography and were evaluated according to the international clinical grading standards for DR established by the American Academy of Ophthalmology^[23].

Definitions

According to the DR Preferred Practice Pattern^[23], DR was clinically classified into two types—non-PDR (NPDR) and PDR; the latter was identified by neovascularization and preretinal or vitreous hemorrhage. To further study the effect of blood pressure-related indicators (SBP and DBP in both the standing and supine positions), obesity-related indicators (BMI and waist circumference), blood glucose-related indicators (Glu0 and HbA1c), blood lipid-related indicators (Chol, Trig, LDL-C, and HDL-C), and renal function related indicators (UA and eGFR) on DR, these indicators were defined according to relevant guidelines or expert consensus. Abnormal blood pressure was defined as SBP ≥ 140 mmHg and/or DBP ≥ 90 mmHg^[24]. Overweight was defined as BMI ≥ 24 kg/m², obesity was defined as BMI ≥ 28 kg/m², and abdominal obesity was defined as waist circumference ≥ 90 cm in men or ≥ 85 cm in women^[25]. Well-controlled blood glucose level was defined as Glu0 ≤ 7.0 mmol/L or HbA1c $< 7\%$, generally controlled blood glucose level was defined as $7\% \leq \text{HbA1c} < 8\%$, and poorly controlled blood glucose level was defined as Glu0 > 7.0 mmol/L or HbA1c $\geq 8\%$ ^[21]. Dyslipidemia was defined as Chol ≥ 6.22 mmol/L, Trig ≥ 2.26 mmol/L, or LDL-C ≥ 4.14 mmol/L and/or HDL-C < 1.04 mmol/L^[26]. Hyperuricemia was defined as fasting serum UA > 420 $\mu\text{mol/L}$ ^[27]. Stages of chronic kidney disease (CKD) were determined by eGFR as follows: CKD stage 1 (G1), eGFR ≥ 90 mL/min; CKD stage 2 (G2), eGFR = 60-89 mL/min; CKD stage 3 (G3), eGFR = 30-59 mL/min; CKD stage 4 (G4), eGFR = 15-29 mL/min; and CKD stage 5 (G5), eGFR < 15 mL/min^[22].

Statistical analysis

Based on fundus photography, participants were divided into groups without DR (WDR group, $n = 826$) or with DR (DR group, $n = 828$). The DR group was further classified into the NPDR ($n = 403$) and PDR ($n = 425$) groups according to severity. All statistical analyses were performed using SPSS (SPSS 20.0, IBM, United States). Differences between the WDR and DR groups or the NPDR and PDR groups were

assessed (Table 1). For continuous variables, Welch's t -test was used for normal distributions, while the Mann-Whitney U test was used for skewed distributions. For categorical variables, the chi-square test was performed. Using DR or PDR as the dependent variable, logistic regression analysis was explored to analyze DR or PDR-related risk factors. Logistic regression analysis was performed using forward selection (likelihood ratio), with $P < 0.05$ as the entry criterion and $P > 0.1$ as the removal criterion. The decision tree method was performed using the chi-squared automatic interaction detector, with 70% participants as the training dataset, and the remaining 30% as the test dataset. The variable assignments used in both logistic regression analysis and decision tree model are presented in Table 2.

RESULTS

Baseline clinical characteristics

The baseline clinical characteristics of all participants are presented in Table 1.

Univariate analysis of DR-related risk factors

As shown in Table 1, the proportion of women, diabetes duration, supine SBP, standing SBP, supine DBP, Chol, BUN, UA, Scr, Glu0, and HbA1c were higher in the DR group than in the WDR group (all P values < 0.05). In contrast, the DR group had a lower BMI, hip circumference, and eGFR than the WDR group (all P values < 0.05).

Univariate analysis of PDR-related risk factors

The proportion of women, supine SBP, supine DBP, Chol, LDL-C, and BUN were significantly higher in the PDR group than in the NPDR group (all P values < 0.05). However, age, hip circumference, and eGFR were significantly lower in the PDR group than in the NPDR group (all P values < 0.05) (Table 1).

Logistic regression analysis of DR-related risk factors

Women had a 1.520 times higher risk of DR [95% confidence interval (CI): 1.218, 1.897] compared to men (Figure 1). Patients with diabetes duration ≥ 10 years had a 2.118-fold higher risk of DR than those with diabetes duration < 10 years (95%CI: 1.661, 2.700). The risk of DR in patients with standing SBP ≥ 140 mmHg was 1.417 times higher than that in patients with standing SBP < 140 mmHg (95%CI: 1.046, 1.919). Compared to patients with supine SBP < 140 mmHg, those with supine SBP ≥ 140 mmHg had a 1.881-fold higher risk of DR (95%CI: 1.399, 2.528). Compared to patients with normal Chol, those with Chol ≥ 6.22 mmol/L had a 1.591 times higher risk of DR (95%CI: 1.104, 2.291). The risk of DR in patients with CKD stages G2, G3, G4, and G5 was 2.206 (95%CI: 1.678, 2.899), 7.860 (95%CI: 4.573, 13.512), 9.693 (95%CI: 3.255, 28.862), and 20.691 (95%CI: 2.540, 168.581) times higher than that in patients with CKD stage G1. In other words, a greater severity of CKD was associated with a higher risk of DR. The risk of DR in patients with $7\% \leq \text{HbA1c} < 8\%$ or $\text{HbA1c} \geq 8\%$ was 1.787 (95%CI: 1.198, 2.664) and 3.073 (95%CI: 2.225, 4.245) times higher than that in patients with $\text{HbA1c} < 7\%$, indicating that worse control of HbA1c was associated with a higher risk of DR.

3 Logistic regression analysis of PDR-related risk factors

Compared to men, women had a 2.161-fold higher risk of progression to PDR (95%CI: 1.615, 2.890) (Figure 2). Compared to patients with diabetes duration < 10 years, patients with diabetes duration ≥ 10 years had a 1.483 times higher risk of PDR (95%CI: 1.099, 2.001). The risk of progression to PDR in patients with CKD stages G3 and G4 was 2.109 (95%CI: 1.362, 3.266) and 2.290 (95%CI: 1.016, 5.165) times higher than that in patients with CKD stage G1.

Decision tree modeling of DR-related risk factors

As shown in Figure 3, the importance of variables in the decision tree model was presented as a root-to-leaf structure, with diabetes duration being the first variable or root node, followed by CKD stage, supine SBP, standing SBP, and BMI, in order of

importance. As presented in Table 3, seven “if-then” rules summarized the path from the root node to each leaf node.

DISCUSSION

In this study, we attempted to reveal the DR-related risk factors in Han population with T2DM in Yunnan province and construct a predictive model for personalized DR risk assessment and early preventive effect.

Studies have reported that various factors modulate the effects of age on DR. Although sporadic cases have been reported, the onset of DR before puberty is extremely rare^[28,29]. Researchers have suggested that patients with diabetes during adolescence are prone to develop serious vascular complications, including DR, compared to patients with diabetes after adolescence. This could be partly due to the characteristics of adolescent patients; for example, patients at this stage tend to be accompanied by dramatic hormone level fluctuations, and most patients present with type 1 diabetes tend to have relatively poor blood glucose self-management ability^[30]. Therefore, this study focused on Han population with T2DM aged ≥ 18 years in Yunnan province to exclude the potential confounding effects of adolescence and type 1 diabetes on the results.

DR has emerged as the leading cause of blindness among 27-75 year olds worldwide^[3,21]. A Chinese meta-analysis reported that the prevalence of DR in patients with T2DM was age related, increasing from 12.55% in adults aged 30-39 years to 20.44% in adults aged 60-69 years and decreasing to 11.22% in those aged ≥ 80 years^[12]. Of the 1654 patients enrolled in this study, 33.49% and 16.56% patients with DR were < 60 years of age ($n = 554$) and ≥ 60 years of age ($n = 274$), respectively. This suggests that the peak of DR prevalence in Han population with T2DM in Yunnan province is concentrated in the population aged < 60 years, which accounts for the majority of the social labor force. In addition, univariate analysis (Table 1) revealed no significant difference in age between the DR and WDR groups ($P = 0.056$). However, overall age was dramatically lower in the PDR group than in the NPDR group ($P < 0.001$). As

shown in Figures 1 and 2, age was not retained in the logistic regression equation. In conclusion, the correlation between age and DR is complex; this association depends on age stratification and may be affected by the degree of vision. This relationship warrants further exploration in future studies.

The relationship between sex and DR is unclear. A study from Germany and Australia based on 120000 samples suggested that women are more likely to develop DR than men^[31]. Similarly, studies in the United Kingdom and Japan have reported that women are more prone to visual impairments than men^[32,33]. Other studies from the United States and India, however, have reported that the men have a higher risk of DR than women^[6,34,35]. In particular, the United Kingdom Prospective Diabetes Study (UKPDS) has proposed that DR progression is associated with the male sex^[36]. Therefore, it is necessary to further explore the correlation between sex and DR. Univariate analysis (Table 1) revealed that the proportion of women was significantly higher in the DR group than in the WDR group ($P = 0.002$). Moreover, the proportion of women was significantly higher in the PDR group than in the NPDR group ($P < 0.001$). Indeed, further logistic regression analysis (Figures 1 and 2) re-emphasizes the importance of female sex. These findings suggest that female sex not only is a risk factor for the development of DR in patients with T2DM but also contributes to the progression of DR to PDR, at least in Yunnan province.

It is well-established that the diabetes duration majorly affects the occurrence and progression of DR^[23,37]. A case-control study in South Korea demonstrated that of 523 patients with T2DM, 44.9% developed DR, and 13.6% developed PDR. The average diabetes duration to mild NPDR, moderate-severe NPDR, and PDR was 14.8, 16.7, and 17.3 years, respectively^[38]. Based on the Chinese population, a meta-analysis indicated that the prevalence of DR in patients newly diagnosed with diabetes and patients with a disease course of ≥ 10 years was 9.00% and 55.52%, respectively^[12]. Univariate analysis (Table 1) revealed that the diabetes duration in the DR group was substantially longer than that in the WDR group ($P < 0.001$). In contrast, the diabetes duration did not differ significantly between the PDR and NPDR groups ($P = 0.221$). However, logistic

regression analysis (Figures 1 and 2) revealed that a diabetes duration ≥ 10 years was an extremely risk for the occurrence and progression of DR. Crucially, the diabetes duration was classified as the root node of the DR decision tree model (Table 3 and Figure 3), emphasizing that the diabetes duration is critical in DR risk assessment.

In addition to the diabetes duration, good glycemic control is considered a key factor for reducing vascular complications of diabetes^[39]. This study focused on two blood-glucose-related indicators, Glu0 and HbA1c. Compared with the transient characteristics of Glu0, HbA1c reflects the overall level of blood glucose control of patients in the prior 2 to 3 months. In this study, univariate analysis (Table 1) indicated that both Glu0 and HbA1c were substantially greater in the DR group than in the WDR group (all $P < 0.05$). As shown in Figure 1, HbA1c but not Glu0 was retained in the logistic regression equation. These findings suggest that poor HbA1c control is associated with a higher risk of DR. In conclusion, compared to Glu0, HbA1c, which reflects long-term glucose control levels, is more relevant for the prevention of DR. However, HbA1c did not negatively affect the progression of DR. Of note, large clinical studies such as the UKPDS^[40,41] and the Diabetes Control and Complications Study^[42] have demonstrated that early and intensive glucose control can reduce the occurrence and progression of diabetic vascular complications, including DR. However, good glycemic control is not equivalent to excessive control. Indeed, extensive evidence indicates that recurrent hypoglycemic episodes caused by excessive strict glycemic control with insulin are associated with the early deterioration of DR, but the underlying mechanisms are unclear^[43-45].

Previous studies have demonstrated that hypertension is linked to the development and severity of DR^[46,47]. Since patients with diabetes are prone to have complications of postural blood pressure changes^[48,49], data on standing and supine blood pressure were collected simultaneously. Univariate analysis (Table 1) revealed that standing or supine SBP and supine DBP were significantly lower in the WDR group than in the DR group (all $P < 0.001$), but there was no obvious difference in standing DBP between the groups ($P = 0.501$). Although supine SBP and supine DBP were lower in the NPDR group than

in the PDR group (all $P < 0.05$), no significant intergroup differences existed in standing SBP and standing DBP (all $P > 0.05$). Further logistic regression analysis (Figures 1 and 2) indicated that SBP had an effect on DR occurrence but not DR progression. In addition, supine SBP ≥ 140 mmHg was the leaf node of the decision tree model, second only to diabetes duration. Our results emphasize the detrimental effects of elevated SBP (especially supine SBP) on DR, suggesting that good blood pressure control is vital for the prevention of DR. Furthermore, the benefits of certain antihypertensive drugs, particularly angiotensin-converting enzyme inhibitors and angiotensin receptor blockers, are not limited to lowering blood pressure^[50,51] and may also benefit DR through neuroprotection^[52,53], increasing insulin sensitivity^[54], anti-inflammatory effects^[55], and inhibiting the blood-eye barrier^[56,57].

To date, the complex link between dyslipidemia and DR has remained controversial^[10,47,58,59]. In this study, the occurrence and development of DR seemed to be more strongly affected by Chol than by Trig, HDL-C, and LDL-C. Univariate analysis (Tables) revealed that Chol was significantly higher in the DR group than in the WDR group ($P < 0.001$) and significantly higher in the PDR group than in the NPDR group ($P = 0.013$). Moreover, as shown in Figures 1 and 2, Chol increased the risk of developing DR in patients with T2DM, but had no remarkable impact on DR progression. A recent meta-analysis revealed that lipid-lowering drugs exerted a protective effect on the progression of DR^[60] but did not affect the deterioration of visual acuity or aggravation of hard exudate. Therefore, further large-scale clinical trials are urgently needed to substantiate the necessity of early application of lipid-lowering drugs in patients with DR.

In parallel with the tremendous rise in the global prevalence of obesity, the prevalence of obesity-related T2DM has also increased annually^[2]. However, the relationship between obesity and DR has not been fully elucidated. Reports from Wisconsin illustrated that obesity (defined by BMI) was not independently implicated in the occurrence or progression of DR in patients with T2DM within 10 years^[61]. Similarly, in the Hoorn study, WHR, but not BMI, was related to the occurrence of

DR^[62]. However, data based on Asian populations have provided the opposite conclusions. For example, based on a sample of 420 Asian patients with T2DM, Man *et al*^[63] reported that BMI was negatively correlated with mild, moderate, and severe DR, but WHR was positively correlated with DR severity. Subsequently, a Korean study demonstrated that higher BMI, increased waist circumference, and higher body fat content (measured by dual-energy X-ray) were notably correlated with a lower risk of DR^[64]. In this study, data regarding relevant body mass indicators were collected, including those on BMI, waist circumference, hip circumference, and WHR. Univariate analysis (Table 1) revealed that BMI and hip circumference in the DR group were remarkably lower than those in the WDR group (all $P < 0.001$), but waist circumference was did not differ across the groups ($P = 0.060$). Although the PDR group had a lower hip circumference than the NPDR group ($P = 0.046$), there were no significant intergroup differences in BMI and waist circumference ($P > 0.05$). In addition, an obvious difference in WHR was not noted among groups (all $P > 0.05$). Although the logistic regression model ultimately did not retain BMI, waist circumference, hip circumference, and WHR (Figures 1 and 2), the decision tree model (Table 3 and Figure 3) supported the protective effect of higher BMI for the assessment of DR risk. In general, although the relationship between these obesity-relevant indicators and DR is yet to be confirmed, it can be conjectured that excessively low BMI is not conducive to protection against DR in the Asian population.

Among BUN, Scr, and UA, eGFR calculated using Scr is the gold standard for CKD staging^[65]. In addition, CKD staging is an essential approach to evaluating the severity of diabetic nephropathy in clinical practice. Univariate analysis (Tables 1) indicated that eGFR was significantly lower in the DR group than in the WDR group ($P < 0.001$) and was significantly lower in the PDR group than in the NPDR group ($P = 0.001$). Furthermore, Figures 1 and 2 highlight the importance of CKD staging in the risk assessment of DR. In this regard, the risk of occurrence and progression of DR increased stepwise with each additional risk level of CKD staging. Notably, in the decision tree model (Table 3 and Figure 3), CKD staging was second only to diabetes duration for DR

risk assessment. In particular, CKD staging is a key indicator of diabetic nephropathy, and the current results also suggest that DR is often comorbid with diabetic nephropathy. This highlights the need to simultaneously screen for diabetic nephropathy in patients with DR.

Differing from traditional statistical methods such as logistic regression analysis, decision trees are successfully employed in the field of medicine with its advantages in solving classification problems, that is, qualitatively judging the possibility of each risk factor at a specific level. Decision trees obtain a set of effective classification rules by systematically learning multiple attributes of the samples with known classification results. When faced with new unknown samples, the appropriate classification or characteristic attributes can be quickly obtained based on the set of rules extracted from the established decision tree^[18-20]. In other words, three basic elements compose the decision tree: root node, internal node, and leaf node. The root node is the main feature attribute in the model, the internal node is the secondary attribute judgment based on the root node, and the leaf node is the final classification outcome of the model.

This study extended traditional statistical analysis by building a DR decision model using machine learning based on the attributes of T2DM samples. In the decision tree model, diabetes duration was demonstrated to primarily affect the occurrence of DR in patients with T2DM, namely, the root node. The extraction rules were interpreted as follows: for patients with diabetes duration < 10 years, if they met the criteria of (1) CKD stage = G3/G4/G5; (2) CKD stage = G2 and BMI < 24 kg/m²; or (3) CKD stage = G1 and standing SBP ≥ 140 mmHg, then DR was prone to occur. In contrast, for patients with diabetes duration ≥ 10 years, if they met the criteria of: (1) Supine SBP < 140 mmHg and CKD stage = G2/G3/G4; or (2) supine SBP ≥ 140 mmHg and CKD stage = G3/G4/G5, then DR was prone to occur. This model may assist clinicians in Yunnan province (particularly primary medical staff who lack relevant DR detection approaches such as ophthalmoscope) to make more effective clinical predictions of DR risk in patients with T2DM. Our decision tree model is simple and intuitive, highlighting its potential for application in clinical practice.

However, there are several limitations that should be noted. First of all, this study did not record in detail the medication of patients, especially ⁶¹ the use of anti-diabetic, lipid-lowering and antihypertensive drugs. Secondly, Yunnan province is located in the western plateau of China, and its climate, cultural and economic conditions are very different from those of plain areas. Therefore, these confounding factors should be included in the future to enhance the integrity and reliability of research conclusions.

CONCLUSION

Female sex, diabetes duration ≥ 10 years, standing or supine SBP ≥ 140 mmHg, Chol ≥ 6.22 mmol/L, deterioration of CKD stage, and HbA1c are key DR-related risk factors in the Han population with T2DM in Yunnan province. The concise and intuitive DR prediction model developed through machine learning in this study could help clinicians quickly predict DR outcomes based on patients' potential risk factors and conduct early individualized interventions.

ARTICLE HIGHLIGHTS

Research background

Yunnan province has a high prevalence of diabetic retinopathy (DR). Accordingly, ² it is of great significance to explore the DR-related factors and to construct an economic and intuitive clinical prediction model.

Research motivation

The research motivation is early intervention using the DR-related risk factors from the perspective of a predictive model to reduce the prevalence ²⁴ of DR in patients with type 2 diabetes mellitus (T2DM).

Research objectives

The research intends to establish a prediction model that allows clinically early prevention and treatment of DR.

Research methods

A total of 1654 Han population with T2DM were recruited in this study and were grouped in the without DR and DR groups. The DR group was further subgrouped according to the severity of DR. Then, univariate analysis, logistic regression analysis, and clinical decision tree models of clinical data were performed.

Research results

Based on the decision tree model constructed in this study, DR classification outcomes were obtained by evaluating diabetes duration followed by stages of chronic kidney disease, supine systolic blood pressure (SBP), standing SBP, and body mass index.

Research conclusions

Personalized interventions for DR-related risk factors based on a decision tree model may potentially reduce the prevalence of DR.

Research perspectives

In this study, patients with T2DM in Western China were taken as samples to analyze the influencing factors of DR and build a clinical prediction model. In the future, it is hoped that the prediction model can produce certain social and economic benefits in clinical practice. In addition, when comparing with other clinical studies on DR, we found some controversies, such as the impact of sex and body mass index on DR, which opened up a new direction for future research.

Figure Legends

Figure 1 Logistic regression analysis of diabetic retinopathy-related risk factors. Female sex, diabetes duration ≥ 10 years, standing systolic blood pressure (SBP) ≥ 140 mmHg, supine SBP ≥ 140 mmHg, cholesterol ≥ 6.22 mmol/L, greater severity of chronic kidney disease, and worse control of hemoglobin A 1c are associated with a higher risk

of diabetic retinopathy. Values are shown using a base 10, logarithmic scale. DR: Diabetic retinopathy; OR: Odds ratio; CI: Confidence interval; BMI: Body mass index; SBP: Systolic blood pressure; Chol: Cholesterol; Trig: Triglyceride; CKD: Chronic kidney disease; HbA1c: Glycated hemoglobin A1c.

Figure 2 Logistic regression analysis of proliferative diabetic retinopathy-related risk factors. Female sex, diabetes duration ≥ 10 years, and chronic kidney disease stage G3 or G4 are risk factors for the progression to proliferative diabetic retinopathy. Values are shown using a base 10, logarithmic scale. Abbreviations: PDR: Proliferative diabetic retinopathy; NPDR: Non-proliferative diabetic retinopathy; OR: Odds ratio; CI: Confidence interval; CKD: Chronic kidney disease.

Figure 3 Training dataset of decision tree model for diabetic retinopathy. Based on the decision tree model constructed in this study, the diabetic retinopathy classification outcomes are obtained by evaluating standing systolic blood pressure (SBP) or body mass index according to the chronic kidney disease (CKD) stage for patients with a diabetes duration < 10 years and the evaluation of CKD stage according to the supine SBP for patients with a diabetes duration ≥ 10 years. WDR: Without diabetic retinopathy; DR: Diabetic retinopathy; CKD: Chronic kidney disease; SBP: Systolic blood pressure; BMI: Body mass index.

Table 1 Univariate analysis of risk factors for diabetic retinopathy or proliferative diabetic retinopathy

Variable	WDR group (<i>n</i> = 826)	DR group (<i>n</i> = 828)	NPDR group (<i>n</i> = 403)	PDR group (<i>n</i> = 425)	<i>P</i> value ¹	<i>P</i> value ²
Age (yr)	53.9 ± 10.3	54.9 ± 10.7	56.2 ± 10.9	53.6 ± 10.3	0.056	< 0.001
Sex (male/female)	487/339	424/404	238/165	186/239	0.002	< 0.001
Diabetes duration (years)	6.9 ± 5.2	11.2 ± 6.8	10.9 ± 7.1	11.4 ± 6.4	< 0.001	0.221
Height (cm)	165.2 ± 8.3	162.5 ± 8.3	164.2 ± 8.5	160.8 ± 7.8	< 0.001	< 0.001
Weight (kg)	68.1 ± 11.7	64.2 ± 11.3	65.7 ± 11.5	62.8 ± 10.9	< 0.001	< 0.001
BMI (kg/m ²)	24.9 ± 3.3	24.3 ± 3.4	24.3 ± 3.3	24.2 ± 3.5	< 0.001	0.880
Waist circumference (cm)	90.8 ± 9.5	89.9 ± 10.0	90.2 ± 9.7	89.6 ± 10.3	0.060	0.399
Hip circumference (cm)	97.6 ± 7.6	95.9 ± 7.4	96.4 ± 7.2	95.4 ± 7.6	< 0.001	0.046
WHR	1.0 ± 0.5	0.9 ± 0.1	0.9 ± 0.1	0.9 ± 0.1	0.284	0.092
Standing SBP (mmHg)	129.0 ± 16.9	136.4 ± 22.9	135.9 ± 22.6	136.9 ± 23.2	< 0.001	0.517
Standing DBP (mmHg)	83.4 ± 11.6	83.8 ± 13.4	83.5 ± 13.0	84.0 ± 13.7	0.501	0.578
Supine SBP (mmHg)	130.1 ± 16.3	141.0 ± 21.3	139.4 ± 20.6	142.4 ± 22.0	< 0.001	0.045
Supine DBP (mmHg)	82.6 ± 11.2	85.0 ± 12.3	83.9 ± 12.0	86.1 ± 12.4	< 0.001	0.012
Glu0 (mmol/L)	9.4 ± 3.3	9.9 ± 3.7	10.0 ± 3.7	9.8 ± 3.7	0.008	0.647
HbA1c (%)	9.1 ± 2.5	9.8 ± 2.4	9.9 ± 2.5	9.7 ± 2.3	< 0.001	0.402

²⁷ Chol (mmol/L)	4.4 ± 1.5	5.0 ± 1.5	4.8 ± 1.5	5.1 ± 1.4	< 0.001	0.013
Trig (mmol/L)	2.8 ± 2.3	2.6 ± 2.4	2.6 ± 2.5	2.6 ± 2.4	0.033	0.856
HDL-C (mmol/L)	1.1 ± 0.3	1.1 ± 0.3	1.1 ± 0.3	1.1 ± 0.4	0.057	0.095
LDL-C (mmol/L)	2.9 ± 0.9	3.0 ± 1.0	2.9 ± 1.0	3.1 ± 1.1	0.101	0.004
BUN (μmol/L)	5.1 ± 1.7	6.6 ± 2.8	6.4 ± 2.8	6.8 ± 2.8	< 0.001	0.021
UA (μmol/L)	347.6 ± 103.4	359.4 ± 102.0	355.4 ± 106.4	363.2 ± 97.6	0.019	0.270
Scr (μmol/L)	69.8 ± 28.3	90.0 ± 61.8	86.0 ± 64.1	93.9 ± 59.4	< 0.001	0.068
eGFR (mL/min)	116.9 ± 37.3	98.3 ± 42.4	103.5 ± 40.9	93.3 ± 43.2	< 0.001	0.001

¹WDR group *vs* DR group.

²NPDR group *vs* PDR group.

¹²WDR: Without diabetic retinopathy; DR: Diabetic retinopathy; NPDR: Non-proliferative diabetic retinopathy; PDR: Proliferative diabetic retinopathy; BMI: Body mass index; WHR: Waist-hip ratio; SBP: Systolic blood pressure; DBP: Diastolic blood pressure; Glu0: Fasting blood glucose; HbA1c: Glycated hemoglobin A1c; Chol: Cholesterol; Trig: Triglyceride; HDL-C: High-density lipoprotein Chol; LDL-C: Low-density lipoprotein Chol; BUN: Blood urea nitrogen; UA: Uric acid; Scr: Serum creatinine; eGFR: Estimated glomerular filtration rate.

Table 2 Variable assignment

Variable	Assignment
Sex	0 = male 1 = female
53 Age	0 = age < 60 years 1 = age ≥ 60 years
Diabetes duration	0 = duration < 10 years 1 = duration ≥ 10 years
1 Blood pressure	0 = normal (SBP < 140 mmHg and/or DBP < 90 mmHg) 1 = abnormal (SBP ≥ 140 mmHg and/or DBP ≥ 90 mmHg)
BMI	0 = normal (BMI < 24 kg/m ²) 1 = overweight (24 kg/m ² ≤ BMI < 28 kg/m ²) 2 = obesity (BMI ≥ 28 kg/m ²)
1 Waist circumference	0 = normal (male < 90 cm, female < 85 cm) 1 = increased (male ≥ 90 cm, female ≥ 85 cm)
Glu0	0 = well controlled (Glu0 ≤ 7.0 mmol/L) 1 = poorly controlled (Glu0 > 7.0 mmol/L)
HbA1c	0 = well controlled (HbA1c < 7%) 1 = generally controlled (7% ≤ HbA1c < 8%) 2 = poorly controlled (HbA1c ≥ 8%)
1 Blood lipids	0 = normal 1 = dyslipidemia (met any of the following criteria: TG ≥ 2.26 mmol/L; TC ≥ 6.26 mmol/L; LDL ≥ 4.14 mmol/L; HDL < 1.04 mmol/L)
UA	0 = normal 1 = hyperuricemia (UA > 420 μmmol/L)
34 CKD stage	0 = G1 (eGFR ≥ 90 mL/min) 1 = G2 (eGFR = 60-89 mL/min)

2 = G3 (eGFR = 30-59 mL/min)

3 = G4 (eGFR = 15-29 mL/min)

4 = G5 (eGFR < 15 mL/min)

15
SBP: Systolic blood pressure; DBP: Diastolic blood pressure; BMI: Body mass index; Glu0: Fasting blood glucose; HbA1c: Glycated hemoglobin A1c; Trig: Triglyceride; Chol: Cholesterol; LDL-C: High-density lipoprotein Chol; HDL-C: high-density lipoprotein Chol; UA: uric acid; CKD: Chronic kidney disease; G: Grade; eGFR: Estimated glomerular filtration rate.

Table 3 “If-then” ² rules extracted from decision tree

Rule	If	Then
R1	Diabetes duration < 10 years, CKD stage = G1, standing SBP < 140 mmHg	Then: WDR
R2	Diabetes duration < 10 years, CKD stage = G1, standing SBP ≥ 140 mmHg	Then: DR
R3	Diabetes duration < 10 years, CKD stage = G2, BMI < 24 kg/m ²	Then: DR
R4	Diabetes duration < 10 years, CKD stage = G2, BMI ≥ 24 kg/m ²	Then: WDR
R5	Diabetes duration < 10 years, CKD stage = G3/G4/G5	Then: DR
R6	Diabetes duration ≥ 10 years, supine SBP < 140 mmHg, CKD staging = G1/G5	Then: WDR
R7	Diabetes duration ≥ 10 years, supine SBP < 140 mmHg, CKD staging = G2/G3/G4	Then: DR
R8	Diabetes duration ≥ 10 years, supine SBP ≥ 140 mmHg, CKD staging = G1/G2	Then: WDR
R9	Diabetes duration ≥ 10 years, supine SBP ≥ 140 mmHg, CKD staging = G3/G4/G5	Then: DR

¹²
R: Rule; CKD: Chronic kidney disease; SBP: Systolic blood pressure; WDR: Without diabetic retinopathy; DR: Diabetic retinopathy.

24%

SIMILARITY INDEX

PRIMARY SOURCES

1	dmsjournal.biomedcentral.com Internet	188 words — 3%
2	www.ncbi.nlm.nih.gov Internet	133 words — 2%
3	"Great Wall International Congress of Cardiology 2020 / Asian Heart Society Congress 2020", Cardiovascular Innovations and Applications, 2020 Crossref	89 words — 1%
4	www.hindawi.com Internet	53 words — 1%
5	worldwidescience.org Internet	49 words — 1%
6	assets.researchsquare.com Internet	42 words — 1%
7	"Abstracts of the IDF Congress in Paris 2003", Diabetologia, 2003 Crossref	39 words — 1%
8	www.scienceopen.com Internet	38 words — 1%

9	D. Müller-Wieland, J. Brandts, M. Verket, N. Marx, K. Schütt. "Chapter 537 Glycaemic Control in Diabetes", Springer Science and Business Media LLC, 2021 Crossref	36 words — 1%
10	zenodo.org Internet	36 words — 1%
11	Yue Zhang, Nana Meng, Zhiping Lv, Hao Li, Yi Qu. "The gene polymorphisms of 1 but not y and 7L2 are associated with diabetic retinopathy in Chinese type 2 diabetes mellitus cases ", Acta Ophthalmologica, 2014 Crossref	35 words — 1%
12	www.researchsquare.com Internet	35 words — 1%
13	www.science.gov Internet	34 words — 1%
14	academic.oup.com Internet	33 words — 1%
15	pure.manchester.ac.uk Internet	32 words — < 1%
16	www.dovepress.com Internet	32 words — < 1%
17	"Abstract", European Journal of Epidemiology, 2009 Crossref	28 words — < 1%
18	lebanesemedicaljournal.org Internet	28 words — < 1%
19	www.researchgate.net	

27 words — < 1%

20 diabetesjournals.org
Internet

25 words — < 1%

21 link.springer.com
Internet

25 words — < 1%

22 journals.plos.org
Internet

23 words — < 1%

23 Suguru Tanaka, Shu-ichi Fujita, Shun Kizawa, Hideaki Morita, Nobukazu Ishizaka. "Association between FGF23, α -Klotho, and Cardiac Abnormalities among Patients with Various Chronic Kidney Disease Stages", PLOS ONE, 2016
Crossref

22 words — < 1%

24 e-dmj.org
Internet

22 words — < 1%

25 Pei-Pei Wu, Qing-Jing Wu, Cai-Xia Liu, Yun-Fei Chen, Qing Xie. "Correlation of neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio and type 2 diabetic retinopathy", International Eye Research, 2020
Crossref

20 words — < 1%

26 sciendo.com
Internet

20 words — < 1%

27 docksci.com
Internet

19 words — < 1%

28 Yuan-Yuan Zhou, Hong-Mei Qiu, Ying Yang, Yuan-Yuan Han. "Analysis of risk factors for carotid intima-media thickness in patients with type 2 diabetes mellitus"

18 words — < 1%

in Western China assessed by logistic regression combined with a decision tree model", Diabetology & Metabolic Syndrome, 2020

Crossref

29 "Abstracts Poster", American Journal of Transplantation, 5/2008 17 words — < 1%
Crossref

30 bio-protocol.org 16 words — < 1%
Internet

31 Ye Fan, Jian-Jun Huang, Chen-Ming Sun, Nan Qiao, Hai-Xia Zhang, Hui Wang, Ran Tao, Ya-Nan Shen, Tong Wang. "Prevalence of dyslipidaemia and risk factors in Chinese coal miners: a cross-sectional survey study", Lipids in Health and Disease, 2017 15 words — < 1%
Crossref

32 www.ijste.org 15 words — < 1%
Internet

33 downloads.hindawi.com 14 words — < 1%
Internet

34 I.J.A. de Bruin, C.E. Wyers, P.C. Souverein, T.P. van Staa et al. "The risk of new fragility fractures in patients with chronic kidney disease and hip fracture—a population-based cohort study in the UK", Osteoporosis International, 2020 12 words — < 1%
Crossref

35 Rana Sorkhabi, Mohammad hossein Ahoor, Amir Ghorbani Haghjo, Elnaz Tabei, Nazli Taheri. "Assessment of tear inflammatory cytokines concentration in patients with diabetes with varying severity of involvement", Experimental Eye Research, 2022 12 words — < 1%
Crossref

36 Xue Tao, Min Jiang, Yumeng Liu, Qi Hu et al. "Predicting three-month fasting blood glucose and glycated hemoglobin of patients with type 2 diabetes based on multiple machine learning algorithms", Research Square Platform LLC, 2022

12 words — < 1%

Crossref Posted Content

37 nccd.org.cn
Internet

12 words — < 1%

38 Peilin Ouyang, You Yiping, Jia Xiaozhou, Yang Liqin. "Risk Prediction of Gestational Diabetes Mellitus In Women With Polycystic Ovary Syndrome Based On A Nomogram Model", Research Square Platform LLC, 2021

11 words — < 1%

Crossref Posted Content

39 discovery.ucl.ac.uk
Internet

11 words — < 1%

40 www.scotland.gov.uk
Internet

11 words — < 1%

41 Aarne Kolonen, Marjatta Sinisalo, Heini Huhtala, Johanna Rimpiläinen et al. " Efficacy of conventional - dose cytarabine, idarubicin and thioguanine versus intermediate - dose cytarabine and idarubicin in the induction treatment of acute myeloid leukemia: Long - term results of the prospective randomized nationwide - 2003 study by the Finnish Leukemia Group ", European Journal of Haematology, 2022

10 words — < 1%

Crossref

42 Anzhela Soloveva, Darya Fedorova, Svetlana Villevalde, Nadezhda Zvartau et al. "Addressing Orthostatic Hypotension in Heart Failure: Pathophysiology,

10 words — < 1%

-
- 43 e-enm.org 10 words — < 1%
Internet
-
- 44 mdpi-res.com 10 words — < 1%
Internet
-
- 45 www.pubfacts.com 10 words — < 1%
Internet
-
- 46 JIAN-GAO FAN. "Fatty liver and the metabolic syndrome among Shanghai adults", Journal of Gastroenterology and Hepatology, 12/2005 9 words — < 1%
Crossref
-
- 47 Tsai, Y-W, M-C Lu, Y-H Lin, Y-C Lee, W-C Li, J-Y Chen, and S-S Chang. "Combined body mass index with high-sensitivity C-reactive protein as independent predictors for chronic kidney disease in a relatively healthy population in Taiwan", European Journal of Clinical Nutrition, 2016. 9 words — < 1%
Crossref
-
- 48 jasn.asnjournals.org 9 words — < 1%
Internet
-
- 49 mafiadoc.com 9 words — < 1%
Internet
-
- 50 oar.a-star.edu.sg 9 words — < 1%
Internet
-
- 51 onlinelibrary.wiley.com 9 words — < 1%
Internet

-
- 52 www.kjim.org 9 words — < 1%
Internet
-
- 53 Žarko Mavrić, Luka Zaputović, Davorka Žagar, Ante Matana, Davor Smokvina. "Usefulness of blood lactate as a predictor of shock development in acute myocardial infarction", *The American Journal of Cardiology*, 1991 9 words — < 1%
Crossref
-
- 54 "ASBMR 2010 Annual Meeting 1001-1264", *Journal of Bone and Mineral Research*, 2010 8 words — < 1%
Crossref
-
- 55 "Abstracts", *Diabetologia*, 1997 8 words — < 1%
Crossref
-
- 56 "Abstracts from the 30th Annual Meeting of the Society of General Internal Medicine", *Journal of General Internal Medicine*, 2007 8 words — < 1%
Crossref
-
- 57 "Poster Session Clinical", *European Journal of Heart Failure*, 2013. 8 words — < 1%
Crossref
-
- 58 H.I. Frohwein, J.H. Lambert, Y.Y. Haimes. "Alternative measures of risk of extreme events in decision trees", *Reliability Engineering & System Safety*, 1999 8 words — < 1%
Crossref
-
- 59 Hang Li, Jie Lin, Yi ying Liu, Yao Liu, Qin Wan. "Association between triglyceride glucose index and carotid atherosclerosis in patients with type 2 diabetes mellitus: a cross-sectional study", *Research Square Platform LLC*, 2021 8 words — < 1%
Crossref Posted Content
-

60 Satoshi Umemura, Hisatomi Arima, Shuji Arima, Kei Asayama et al. "The Japanese Society of Hypertension Guidelines for the Management of Hypertension (JSH 2019)", Hypertension Research, 2019

8 words — < 1%

Crossref

61 Shaopeng Yang, Yixiong Yuan, Yanping Chen, Shiran Zhang et al. "Metabolic signature of the ganglion cell inner plexiform layer thickness and the risks of mortality and morbidity: a population-based study in UK Biobank", Cold Spring Harbor Laboratory, 2022

8 words — < 1%

Crossref Posted Content

62 dlisv03.media.osaka-cu.ac.jp

Internet

8 words — < 1%

63 "2017 Annual Meeting of the American Society for Bone and Mineral Research, Colorado Convention Center, Denver, CO, USA - September 8-11, 2017", Journal of Bone and Mineral Research, 2017

7 words — < 1%

Crossref

64 Jae-Seung Yun, Tae-Seok Lim, Seon-Ah Cha, Yu-Bae Ahn et al. "Clinical Course and Risk Factors of Diabetic Retinopathy in Patients with Type 2 Diabetes Mellitus in Korea", Diabetes & Metabolism Journal, 2016

7 words — < 1%

Crossref

65 Zhenzhen Liu, Mingxi Shao, Jun Ren, Yichao Qiu, Shengjie Li, Wenjun Cao. "Association Between Increased Lipid Profiles and Risk of Diabetic Retinopathy in a Population-Based Case-Control Study", Journal of Inflammation Research, 2022

7 words — < 1%

Crossref

66 "Abstracts", Diabetologia, 2005

6 words — < 1%

Crossref

67 Jia Pan, Bo Yang, Zihang Wang, Lei Tang, Peng Jia, Shujuan Yang, Honglian Zeng. "Association between Triglyceride-glucose index and carotid artery plaque in young and middle-aged occupational population", Research Square Platform LLC, 2022

6 words — < 1%

Crossref Posted Content

68 Shijie Sun, Minglong Li, Jie Zhou, Zhibo Gai, Haiyan Shi, Qing Zhao, Jun Tian. "Cystatin C predicts diabetic retinopathy in Chinese patients with type 2 diabetes", International Journal of Diabetes in Developing Countries, 2015

6 words — < 1%

Crossref

69 Xiaokang Dong, Honglei Zhang, Fang Wang, Xiaotian Liu et al. "Epidemiology and prevalence of hyperuricemia among men and women in Chinese rural population: The Henan Rural Cohort Study", Modern Rheumatology, 2019

6 words — < 1%

Crossref

70 Xiaonuo Li, Shiyi Yi, Andrew B. Cundy, Weiping Chen. "Sustainable decision-making for contaminated site risk management: A decision tree model using machine learning algorithms", Journal of Cleaner Production, 2022

6 words — < 1%

Crossref

71 Y. Ding, Q. Ge, H. Qu, Z. Feng, J. Long, Q. Wei, Q. Zhou, R. Wu, L. Yao, H. Deng. "Increased serum periostin concentrations are associated with the presence of diabetic retinopathy in patients with type 2 diabetes mellitus", Journal of Endocrinological Investigation, 2018

6 words — < 1%

Crossref

72 livrepository.liverpool.ac.uk

Internet

6 words — < 1%

EXCLUDE QUOTES OFF
EXCLUDE BIBLIOGRAPHY OFF

EXCLUDE SOURCES OFF
EXCLUDE MATCHES OFF