

## RESEARCH ARTICLE

# Biochemical testing for the diagnosis of Wilson's disease: A systematic review

Hafiz Muhammad Salman<sup>1</sup> | Mahwish Amin<sup>2</sup> | Javaria Syed<sup>3</sup> | Zouina Sarfraz<sup>4</sup>  | Azza Sarfraz<sup>5</sup>  | Muzna Sarfraz<sup>6</sup> | Maria Jose Farfán Bajaña<sup>7,8</sup> | Miguel Felix<sup>7,8</sup> | Ivan Cherez-Ojeda<sup>7,8</sup>

<sup>1</sup>Services Institute of Medical Sciences, Lahore, Pakistan

<sup>2</sup>Sir Ganga Ram Hospital, Lahore, Punjab, Pakistan

<sup>3</sup>Sargodha Medical College, Sargodha, Pakistan

<sup>4</sup>Fatima Jinnah Medical University, Lahore, Pakistan

<sup>5</sup>The Aga Khan University, Karachi, Pakistan

<sup>6</sup>King Edward Medical University, Lahore, Pakistan

<sup>7</sup>Universidad Espíritu Santo, Samborondón, Ecuador

<sup>8</sup>Respiralab, Respiralab Research Group, Guayaquil, Ecuador

## Correspondence

Zouina Sarfraz, Fatima Jinnah Medical University, Lahore 54000, Pakistan.  
Email: zouinasarfraz@gmail.com

Ivan Cherez-Ojeda, Universidad Espíritu Santo, Km. 2.5 Vía La Puntilla Samborondón 0901-952, Ecuador.  
Email: ivancherez@gmail.com

## Abstract

**Background:** Wilson's disease (WD) is a rare inherited disorder that leads to copper accumulation in the liver, brain, and other organs. WD is prevalent worldwide, with an occurrence of 1 per 30,000 live births. Currently, there is no gold standard diagnostic test for WD. The objective of this systematic review is to determine the diagnostic accuracy for WD of three biochemical tests, namely hepatic copper, 24-hour urinary copper, and ceruloplasmin using the Leipzig criteria.

**Methods:** Adhering to PRISMA guidelines, databases including PubMed/MEDLINE, CINAHL Plus, Web of Science, and Cochrane were searched. Studies that comprised of confirmed or suspected WD along with normal populations were included with adult and pediatric group. The sensitivity, specificity, negative predictive value and positive predictive value were computed using RevMan 5.4.

**Results:** Nine studies were included. The best practice evidence for 24-hour urinary copper test ranged from a cutoff value of 0.64–1.6  $\mu\text{mol}/24\text{ h}$  ( $N = 268$ ; sensitivity = 75.6%, specificity = 98.3%). Hepatic copper test was optimally cutoff based on the ROC curve analysis at 1.2  $\mu\text{mol}/\text{g}$  yielding a power of 96.4% sensitivity and 95.4% specificity ( $N = 1,150$ ); however, the tried and tested 4  $\mu\text{mol}/\text{g}$  cutoff, with 99.4% sensitivity and 96.1% specificity, is more widely accepted. The ceruloplasmin test cutoff value was found to be ranging from 0.14 to 0.2 g/L ( $N = 4,281$ ; sensitivity = 77.1%–99%, specificity = 55.9%–82.8%).

**Conclusion:** This paper provides a large-scale analysis of current evidence pertaining to the biochemical diagnosis of WD employing the Leipzig criteria. The laboratory values are typically based on specific subgroups based on age, ethnicity, and clinical subgroups. The findings of this systematic review must be used with caution, given the over- or under-estimation of the index tests.

## KEYWORDS

ceruloplasmin, hepatic copper, hepatolenticular degeneration, liver copper, urinary copper, Wilson's disease

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Clinical Laboratory Analysis* published by Wiley Periodicals LLC.

## 1 | INTRODUCTION

Wilson's disease (WD) was first described in 1912 by Samuel Wilson as an autosomal recessive metabolic disorder occurring due to mutations of the ATP7B gene.<sup>1</sup> It is a rare inherited disorder that leads to copper accumulation in the liver, brain, and other vital organs.<sup>2</sup> WD is found worldwide, with an estimated prevalence of 1 per 30,000 live births across populations,<sup>3</sup> although the data obtained by molecular sequencing from the United Kingdom suggest higher prevalence of 1 per 7,021 live births.<sup>4</sup> A large proportion of patients are diagnosed between the ages of 5 and 35 years, but it may affect the older population as well.<sup>1,2</sup> At present, there is no gold standard diagnostic test for Wilson's disease, where only the measurements of liver copper content are being used to improve the diagnostic accuracy.<sup>5</sup> Owing to the nonspecific clinical features of Wilson's disease, the battery of laboratory and clinical tests for diagnosis is oftentimes delayed.<sup>6</sup> Ultimately, this may affect the clinical outcomes and has implications for members in the family tree when considering late or missed diagnosis earlier during the disease course.<sup>7,8</sup> With the Leipzig criteria, the shortcoming of no gold standard diagnostic test may be overcome by promoting the standardization of diagnosis and treatment of disease.<sup>9</sup>

The objective of this systematic review is to determine the diagnostic accuracy, that is, sensitivity and specificity of biochemical tests, including hepatic copper, 24-hour urinary copper content, and ceruloplasmin. These laboratory markers are ideally tested in suspected patient and control groups. With inconsistent gold standard testing for the diagnosis of Wilson's disease, the Leipzig criteria are used as the standard for this investigation. The key purpose is to scale the benefits of these tests for patients under an index of suspicion or whether the test should be used at large.

## 2 | MATERIALS AND METHODS

In accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement 2020,<sup>10</sup> observational studies (retrospective/prospective cohorts and case controls) that tested the diagnostic accuracy of any or all three index tests in the context of diagnosing WD were included. Databases including the following were searched: PubMed/MEDLINE, CINAHL Plus, Web of Science, and Cochrane. The date of the last database search was October 20, 2021. There were no language restrictions. A combination of the following keywords was used: Wilson's disease, Wilson's disease, ceruloplasmin, urinary copper, hepatic copper, liver copper, and hepatolenticular degeneration. The studies included participants with confirmed or suspected WD and also comprised of normal population and heterozygotes in cases where genetic testing was used. In case the study evaluated the index test in the normal population with the Leipzig criteria but did not include a WD comparator group, it was omitted. The target population was the pediatric and adult population with

suspected WD as assessed by the Leipzig criteria, which was essential for inclusion. Studies that did not use the Leipzig criteria and did not define WD were excluded. The three index tests were ceruloplasmin, liver copper content, and urinary copper that were evaluated for diagnosing WD. The cutoff thresholds for each of the index test are provided in Figure 1. The clinical reference standard to the diagnosis of WD is outlined in the Leipzig criteria below.

Two pairs of authors (AS-ZS and HMS-JS) independently extracted data onto a google spreadsheet, and any discrepancies were resolved by active discussion. The data were entered as author and year, participant and methodology, clinical characteristics, index test details, the reference standard used in the study, and biases of the included studies. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated using ReviewManger 5.4 (RevMan 5.4, Cochrane), which are tabulated for all three index tests in Tables 1–3. The data were entered into the software as the number of TP/ participants with WD (TP + FN). The higher the sensitivity for a particular cutoff, the better the test was considered in correctly identifying individuals with WD. The specificity was calculated as the number of TN/ participants without WD (TN + FP). The higher the specificity of the cutoff value, the better the diagnostic test in identifying individuals who do not have WD. PPV is the number of TP/ numbers of TP + FP. PPV measures the individuals with positive tests who have WD. NPV is the number of TN/numbers of TN +FN. NPV measures the number of people with a negative test result that do not have WD. The methodology encompassed a diagnostic analytical technique to determine the diagnostic cut-offs of the biochemical tests.

## 3 | RESULTS

The PRISMA flowchart is depicted in Figure 2. In total, 13,783 studies were identified from the enlisted databases. On removing 6,436 duplicates, 7,347 studies were screened. Overall, 21 studies were retrieved and thereby assessed for eligibility. Finally, 9 studies were included in this analysis.

### 3.1 | Characteristics of included studies

The characteristics of the included studies are listed in Table 4. Only 5 of the 9 studies provided data for ceruloplasmin. Overall, while 6 of the 9 studies were case controls, with one each being a cross-sectional, a prospective cohort, and a genetic cohort study, they were well-designed and had robust inclusion criteria along with a battery of index testing. Individualistic differences of index tests, patient populations, and evaluations are listed (Table 4). A proportion of the sample size presented mild hepatic WD, whereas neurological cases and asymptomatic WD presentations were also included.

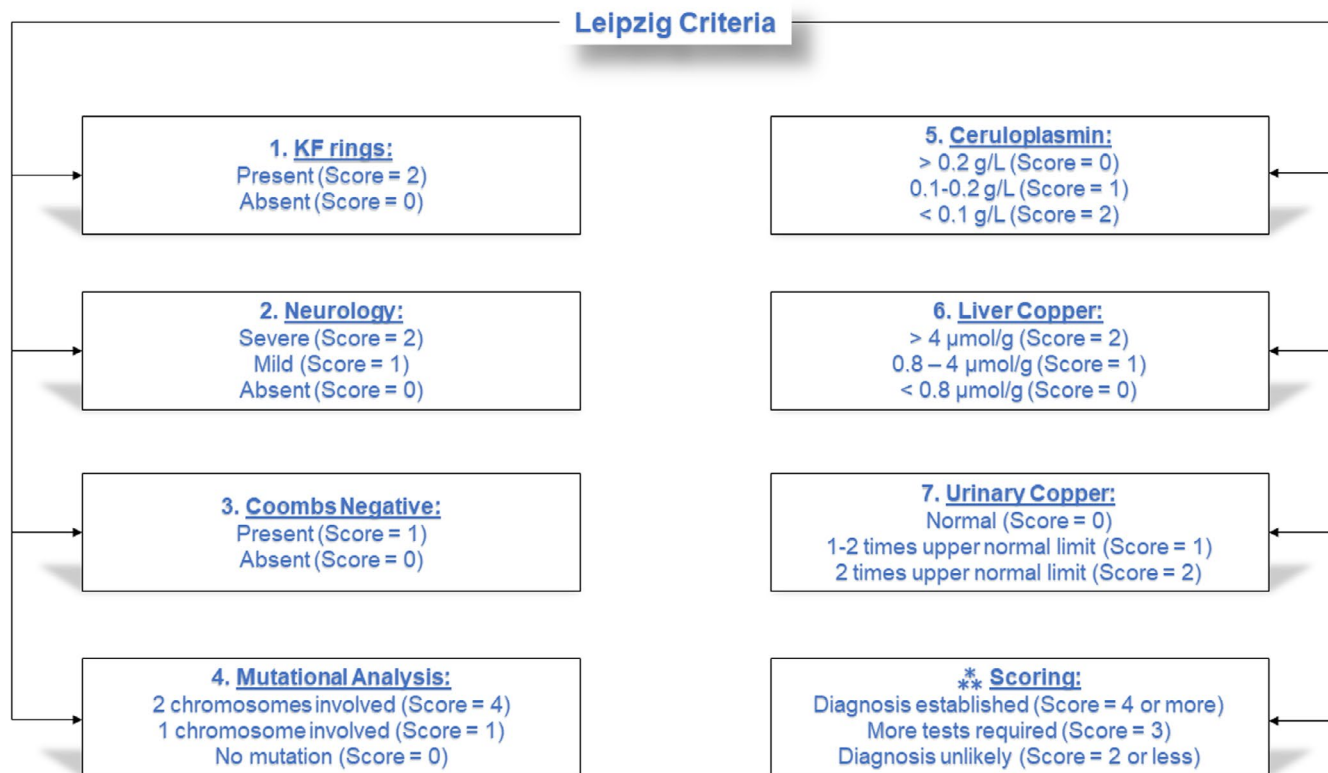


FIGURE 1 Diagnostic algorithm for Wilson's disease based on the Leipzig score

TABLE 1 Ceruloplasmin diagnostic accuracy (sensitivity, specificity, PPV, and NPV using 95% CI)

Author (Year)	Threshold	Sensitivity vs. Specificity (95% CI)	PPV vs. NPV (95% CI)
Mak et al. (2008) <sup>15</sup>	0.1 g/L	78.9% (CI: 66.1–88.6) vs. 100% (CI: 97.3–100)	100% (CI: NE) vs. 91.9% (CI: 87.3–94.9)
Nicastro et al. (2010) <sup>11</sup>	0.1 g/L	65% (CI: 66.1–88.6) vs. 96.6% (CI: 88.1–99.6)	92.9% (CI: 76.6–98.1) vs. 80% (CI: 72.3–86)
Sezer et al. (2014) <sup>12</sup>	0.2 g/L	77.1% (CI: 59.9–89.6) vs. 65.9% (CI: 49.4–79.9)	65.9% (CI: 54.9–75.4) vs. 77.1% (CI: 63.9–86.6)
Xu et al. (2018) <sup>13</sup>	0.2 g/L	99% (CI: 97.1–99.8) vs. 80.9% (CI: 79.6–82.2)	29.1% (CI: 27.8–30.5) vs. 99.9% (CI: 99.7–100)
Merle et al. (2009) <sup>14</sup>	0.19 g/L	93.6% (CI: 87.3–97.4) vs. 58.8% (CI: 44.2–72.4)	83.1% (CI: 77.9–87.2) vs. 81% (CI: 66.9–90.1)

Abbreviations: CI, confidence interval; NE, not estimated; NPV, negative predictive value; PPV, positive predictive value.

TABLE 2 24-hour urinary copper test diagnostic accuracy (sensitivity, specificity, PPV, and NPV using 95% CI)

Author (Year)	Threshold	Sensitivity vs. Specificity (95% CI)	PPV vs. NPV (95% CI)
Nicastro et al. (2010) <sup>11</sup>	0.64 $\mu\text{mol}/24\text{ h}$	78.9% (CI: 62.7–90.5) vs. 87.9% (CI: 76.7–95)	81.1% (CI: 67.7–89.7) vs. 86.4% (CI: 77.4–92.2)
Nicastro et al. (2010a) <sup>11</sup>	1.6 $\mu\text{mol}/24\text{ h}$	65.8% (CI: 48.7–80.4) vs. 98.3% (CI: 90.8–100)	96.2 (CI: 77.9–99.4) vs. 81.4 (CI: 73.8–87.2)
Lu et al. (2010) <sup>16</sup>	1.6 $\mu\text{mol}/24\text{ h}$	50% (CI: 29.9–70.1) vs. 97.1% (CI: 89.8–99.6)	86.7% (CI: 61.1–96.4) vs. 83.5% (CI: 77.5–88.2)
Sezer et al. (2014) <sup>12</sup>	1.6 $\mu\text{mol}/24\text{ h}$	80% (CI: 63.1–91.6) vs. 75.6% (CI: 59.7–87.6)	73.7% (CI: 61.4–83.1) vs. 81.6% (CI: 69.1–86.4)

Abbreviations: CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.

### 3.2 | Ceruloplasmin index test

Five of the 9 studies evaluated the threshold of ceruloplasmin across 4,281 individuals, among which 541 had WD. The cutoffs were defined by the Leipzig criteria (0.1 g/L and 0.2 g/L; 0.19 g/L) (Mak et al., 2008, Nicastro et al.<sup>11</sup> 2010, Sezer et al.<sup>12</sup> 2014, Xu et al.<sup>13</sup> 2018, Merle et al.<sup>14</sup> 2009) (Table 1). In these five studies, the optimal

cutoff for the ceruloplasmin index test was determined to be between 0.14 and 0.2 g/L. The studies were well-designed, clearly defined WD, and the details of the laboratory cutoffs were conducted using receiver operating characteristic (ROC) curve analysis. It may be noted that ceruloplasmin levels are lower in the neonatal age group; however, the levels rise in women who are currently pregnant or taking oral contraception and those undergoing acute

TABLE 3 Hepatic copper test diagnostic accuracy (sensitivity, specificity, PPV, and NPV using 95% CI)

Author (Year)	Threshold	Sensitivity vs. Specificity (95% CI)	PPV vs. NPV (95% CI)
Ferenci et al. (2005) <sup>17</sup>	>4 $\mu\text{mol/g}$	88.3% (CI: 75.2–89.7) vs. 98.6% (CI: 96.1–99.7)	96.9% (CI: 91.1–99) vs. 91.9% (CI: 88.3–94.5)
Ferenci et al. (2005a) <sup>17</sup>	>1.2 $\mu\text{mol/g}$	96.5% (CI: 91.3–99) vs. 95.4% (CI: 91.8–97.8)	91.7% (CI: 85.7–95.3) vs. 98.1% (CI: 95.2–99.3)
Nicastro et al. (2010) <sup>11</sup>	>4 $\mu\text{mol/g}$	93.3% (CI = 77.9–99.2) vs. 52.2% (CI = 37–67.1)	56% (CI: 48.1–63.6) vs. 92.3% (CI: 75.4–97.9)
Sezer et al. (2014) <sup>12</sup>	>4 $\mu\text{mol/g}$	65.7% (CI = 47.8–80.9) vs. 75.6% (CI = 59.7–87.6)	69.7% (CI: 56.1–80.6) vs. 72.1% (CI: 61.3–80.8)
Yang et al. (2015) <sup>18</sup>	>4 $\mu\text{mol/g}$	94.4% (CI = 89.9–97.3) vs. 96.8% (CI = 94.7–98.2)	91.8% (CI: 87.2–94.9) vs. 97.8% (CI: 96.1–97.5)

Abbreviations: CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.

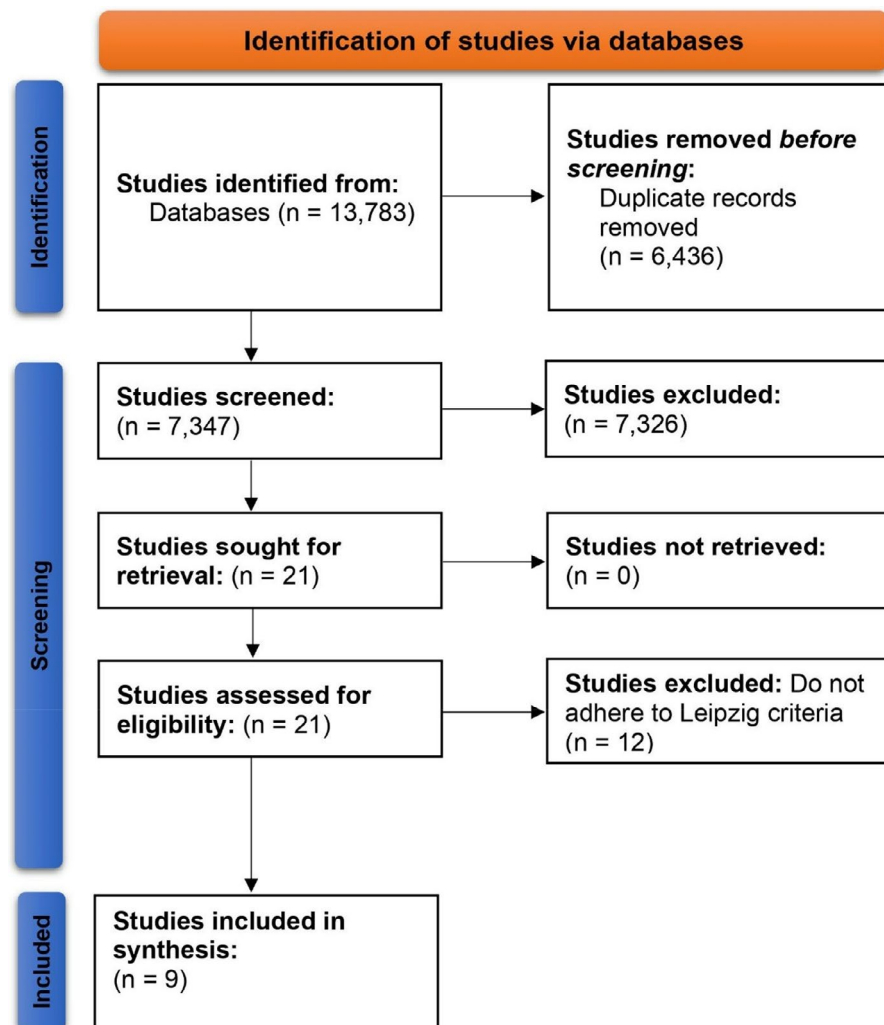


FIGURE 2 PRISMA flowchart

inflammation. Nicastro et al.<sup>11</sup> (2010) and Sezer et al.<sup>12</sup> (2014) conducted their studies in the pediatric population; Mak et al.<sup>15</sup> (2008) and Xu et al.<sup>13</sup> (2018) used a mixed pediatric and adult population, and Merle et al.<sup>14</sup> (2009) only enrolled the adult population. Because ceruloplasmin is synthesized in the liver, the levels are lower in other causes of chronic liver disease, enteropathies, and nephrotic syndrome as compared to WD. Literature suggests that individuals undergoing chelation therapy have lower levels of ceruloplasmin levels; however, Merle et al.<sup>14</sup> present differing results where 65% of the sample was on penicillamine therapy but the minute differences in assay technology may have led to this effect. It is imperative to account for the different methods of the analyte—ceruloplasmin,

which may vary in terms of reference range, precision, and bias. Therefore, the cutoff values identified as 0.14–0.2 g/L serve as an essential method that is paramount when the same sample is run using the same analyte with a different method, but will possibly lead to different results as seen in the Merle study.<sup>14</sup>

### 3.3 | 24-Hour urinary copper test

At present, there is limited evidence pertaining to the adult cutoff, with data enlisted by the European Association for the Study of the Liver (EASL) of 1.6  $\mu\text{mol}/24\text{ h}$ . However, this systematic review

TABLE 4 Characteristics of included studies

Author (Year)	Title	Study type	Sample characteristics	Index laboratory testing	Biases
Ferenci et al. (2005) <sup>17</sup>	Diagnostic value of quantitative hepatic copper determination in patients with Wilson's disease	Case control*	Individuals with WD, of neuropsychiatric, hepatic, or asymptomatic type, controls without WD and participants with other hepatic pathologies	Hepatic copper was measured by atomic absorption spectroscopy	Given the non-randomized enrollment and observational nature of the study, potential biases may have existed
Lu et al. (2010) <sup>16</sup>	The reassessment of the diagnostic value of 24-hour urinary copper excretion in children with Wilson's disease	Case control*	The individuals had unknown hepatic pathologies	24-hour urinary copper was measured by ICP mass spectrometry	Non-randomization, selection bias of the case-control design may have led to bias
Mak et al. (2008) <sup>15</sup>	Diagnostic accuracy of serum ceruloplasmin in Wilson's disease: determination of sensitivity and specificity by ROC curve analysis among ATP7B-genotyped subjects	Case control*	Individuals with WD of neuropsychiatric, hepatic, or asymptomatic type, groups with no diagnosis of hepatic or neurological deficit, and normal controls	Ceruloplasmin with nephelometry Beckman Coulter IMMAGE	The nature of clinical reference standards in addition to non-randomization and low acceptability of results with small sample size may lead to the risk of bias
Merle et al. (2009) <sup>14</sup>	Serum ceruloplasmin oxidase activity is a sensitive and highly specific diagnostic marker for Wilson's disease	Case control*	Individuals with WD, neurological or hepatic, with alternative hepatic pathologies, or normal controls	Ceruloplasmin nephelometry Dade Behring	Sample selection, non-randomization, and case-control amplify the risk of bias
Nicastro et al. (2010) <sup>11</sup>	Re-evaluation of the diagnostic criteria for Wilson's disease in children with mild liver disease	Case control*	Confirmed WD patients either with asymptomatic family screening or hepatic etiology, with alternative hepatic pathologies, or normal controls	Hepatic copper with flame absorption spectrophotometry; urine copper using flame absorption spectrophotometry; ceruloplasmin using radial immunodiffusion NOR-Partigen Behring	Case-control study type increases the risk of non-randomization and less diverse sample set
Sezer et al. (2014) <sup>12</sup>	Is it necessary to re-evaluate diagnostic criteria for Wilson's disease in children?	Case control*	Patients with hepatic WD, alternative hepatic pathologies, or normal controls	Ceruloplasmin by immunoturbidimetry Roche Modular; urine copper with atomic absorption spectrophotometry AA-6701F Shimadzu; and hepatic copper by atomic absorption spectrophotometry AA-6701F Shimadzu	Selection bias, non-randomization bias may increase the risk of bias or confounders

(Continues)

TABLE 4 (Continued)

Author (Year)	Title	Study type	Sample characteristics	Index laboratory testing	Biases
Xu et al. (2018) <sup>13</sup>	The optimal threshold of serum ceruloplasmin in the diagnosis of Wilson's disease: A large hospital-based study	Cross-sectional	Patients that underwent ceruloplasmin analysis were eligible to be included, and the tests/records were noted in a hospital center	Beckman Coulter Image; ceruloplasmin using nephelometry	The cross-sectional study design leads to increase the risk of selection bias
Yang et al. (2015) <sup>18</sup>	Prospective evaluation of the diagnostic accuracy of hepatic copper content, as determined using the entire core of a liver biopsy sample	Prospective cohort	Patients with suspected hepatic WD; family member of people with confirmed WD, or those that had alternative hepatic pathologies	Hepatic copper using atomic absorption spectrophotometry Beijing Purkinje General Instruments	The study had the minimal risk of bias, and the population was most reflective of clinical practice
Zarina et al. (2019) <sup>21</sup>	Association of Variants in the CP, ATOX1, and COMMD1 genes with Wilson's disease symptoms in Latvia	Genetic prospective cohort	Patients with WD: asymptomatic, hepatic, neurological/psychiatric, and neurological/hepatic	Direct sequences of the ATOX1, COMMD1, and CP genes; direct DNA sequencing of the ATP7B gene	The study had the minimal risk of bias other than the relatively less sample size as compared to a large genetic cohort study

Note: \*All patients were being tested for WD, and the Leipzig criteria were the standard of reference.

presents three studies that enlist the 24-hour urinary copper cut-offs of children with WD-associated CLD (Lu et al.<sup>16</sup> 2010, Nicastro et al.<sup>11</sup> 2010, Sezer et al.<sup>12</sup> 2014) (Table 2). Lu and colleagues utilized 24-hour urinary copper was measured by ICP mass spectrometry using a Chinese population,<sup>16</sup> whereas Nicastro/Sezer and et al.<sup>11,12</sup> used an atomic absorption spectroscopy in an Italian and Turkish population, respectively (Table 2). Both these studies used a clear criterion for diagnosing WD, and there were gender- and age-matched controls. Wherever differences in cutoffs were present, they may have occurred due to slight differences in methods and ethnicity.

### 3.4 | Hepatic copper

On considering the cutoffs for copper, four studies in total were eligible, and differences in age groups, methods, and index test analysis were presented (Table 3). Ferenci and colleagues clearly defined the WD criteria and employed a pediatric and adult population, and adequate laboratory methods were utilized to evaluate hepatic copper by using atomic absorption spectroscopy.<sup>17</sup> The optimal cutoff based on the ROC curve analysis was 1.2  $\mu\text{mol/g}$ , giving it a power of 96.4% sensitivity and 95.4% specificity. This is in opposition to the higher cutoff used in the Leipzig criteria of 4  $\mu\text{mol/g}$ . However, the Leipzig criteria for hepatic copper were originally based on a sample of 7 individuals with WD as posited by Ferenci and colleagues.<sup>17</sup> When noting the study by Yang and colleagues, which employed a Chinese sample set, the authors supported the 4  $\mu\text{mol/g}$  cutoff, with 99.4% sensitivity and 96.1% specificity.<sup>18</sup> Yang and authors had a larger sample size as compared to the former studies and clearly defined the criteria and methodology of the adult sample included.<sup>18</sup> Minute differences may exist because distinct methodologies may also reflect ethnic and racial differences in alleles causing disease in the populations. Both Sezer and Nicastro tested hepatic copper using different methodologies of flame atomic absorption spectroscopy in the pediatric population with hepatic disease in Turkey and Italy, respectively.<sup>11,12</sup> While Nicastro did not undertake a ROC analysis, the authors used a 4  $\mu\text{mol/g}$  cutoff based on the Leipzig criteria, yielding a sensitivity of 65% and specificity of 77%.<sup>11</sup> On the contrary, Sezer noted that when the cutoff was decreased to a threshold of 1.5  $\mu\text{mol/g}$ , the sensitivity increased to 91.4%, but the specificity decreased to 65.8%.<sup>12</sup>

## 4 | DISCUSSION

This systematic review assesses the clinical, biochemical, immunological, and genetic tests that are included in the heterogeneous process of diagnosing WD. It may be stated that the variability in optimal cutoff values according to ROC curve analysis depends on the applied biochemical tests methodology and the age of participants as claimed by the evidence provided by the studies in this review. Notably, EASL guidelines supply a narrative expert-base review about current evidence supporting the original Leipzig criteria.

Cauza et al.<sup>19</sup> conducted an assessment of 17 patients that had ceruloplasmin levels less than 20 mg/dl. While only one asymptomatic patient had WD with no neurological signs of Kayser-Fleischer rings, the other 16 patients were considered to be heterozygous carriers of the WD gene. The positive predictive value of low ceruloplasmin was merely 5.9%. Despite considerations regarding the measurement of hepatic copper content as the gold standard for diagnosing WD, Ferenci and colleagues, in one of the largest studies in this area, highlight the occurrence of false-negative results due to sampling errors or differences in hepatic copper distribution.<sup>17</sup>

Nine studies, including a total of 5,762 participants (which comprised of 1,163 individuals with Wilson's disease) were assessed. Three studies involved the pediatric population only, one comprised of adults only, and the other five studies involved both adult and pediatric populations. Two studies assessed individuals with hepatic signs only, and the other seven studies assessed the combination of neurological and hepatic signs and symptoms of WD. There were variable methodological qualities of the studies, the risk of bias due to the study types was documented, and the key differences in assays and cutoff values for the diagnostic tests are enlisted in Tables 1–4.

When noting the ceruloplasmin index test, a total of five studies assessed the thresholds, pooling in 4,281 individuals with 541 confirmed WD patients. The all-around scalable cutoff value was found to be ranging from 0.14 to 0.2 g/L. When using a cutoff of 0.2 g/L as per the Leipzig criteria, a sensitivity of 77.1%–99% was achieved and the specificity ranged from 55.9% to 82.8%. On noting the 0.1 g/L cutoff as per the Leipzig criteria, the sensitivity ranged from 65% to 78.9%, whereas the specificity ranged from 96.6% to 100%. The best practice guidelines resonate that a highly sensitive test leads to fewer false-negative results, and fewer WD cases are missed, whereas the specificity of the ceruloplasmin test means that there are fewer false-positive results. The most scalable test result has maximum sensitivity and specificity. While the current literature states that the ceruloplasmin level for WD will probably be below 0.1 g/L, our synthesis identifies,<sup>20</sup> based on the data obtained by five studies, the optimal cutoff for the ceruloplasmin index test was determined to be between 0.14 and 0.2 g/L.

The 24-hour urinary copper test was evaluated based on three studies, and there were varying thresholds using a participant size of 268, with 101 confirmed WD patients. While EASL posts 1.6  $\mu\text{mol}/24\text{ h}$  as the cutoff, we found that the best practice evidence suggests a cutoff ranging from 0.64 to 1.6  $\mu\text{mol}/24\text{ h}$  based on the Leipzig criteria, which achieves a sensitivity vs. specificity of 50% vs. 80% and 75.6% vs. 98.3%, respectively.

The hepatic copper test assessed a total of 1150 individuals with 367 WD patients. The hepatic cutoff in the Leipzig criteria achieved a sensitivity of 65.7%–94.4% and specificity of 52.2%–98.6%.

Overall, the cutoffs for ceruloplasmin, 24-hour urinary copper, and hepatic copper test are reliant on the methodologies adopted by the testing facility and do require validation in the population when the index tests are used at large. While we provide cutoffs

of the three tests for WD, we find that cutoffs are typically based on the specific subgroups based on ethnicity, age, and clinical subgroups.

While the findings of this systematic review may be used by the readers, and despite the included studies employing the Leipzig WD criteria, some of the included studies had limited details about the methodologies and the calculations, despite mentioning the index tests being appraised. While the current literature addresses the possible gold standards to diagnose WD, there is a lack of disease definition and there is a poorly defined index test methodology. Moreover, while ceruloplasmin levels are helpful in diagnosing WD, the feasibility and cost-effectiveness of utilizing serum ceruloplasmin in presymptomatic WD may be low.<sup>19</sup> With these limitations in mind, we believe that the methodological weaknesses of the included studies may have resulted in overestimation of the accuracy of the index tests, despite still offering pertinent information to medical communities worldwide.

## ACKNOWLEDGEMENTS

None.

## CONFLICT OF INTEREST

None.

## DATA AVAILABILITY STATEMENT

All data used and acquired for this study are available online.

## ORCID

Zouina Sarfraz  <https://orcid.org/0000-0002-5132-7455>

Azza Sarfraz  <https://orcid.org/0000-0001-8206-5745>

## REFERENCES

1. Bandmann O, Weiss KH, Kaler SG. Wilson's disease and other neurological copper disorders. *Lancet Neurol*. 2015;14(1):103-113.
2. Ryan A, Nevitt SJ, Tuohy O, Cook P. Biomarkers for diagnosis of Wilson's disease. *Cochrane Database Syst Rev*. 2019;2019(11):CD012267.
3. Huster D. Wilson disease. *Best Pract Res Clin Gastroenterol*. 2010;24(5):531-539. doi:10.1016/J.BPG.2010.07.014
4. Abid K, Bari YA, Younas M, Tahir Javaid S, Imran A. Progress of COVID-19 epidemic in Pakistan. *Asia-Pacific J Public Heal*. Published online. 2020;32(4):154-156. doi:10.1177/1010539520927259
5. Stremmel W, Merle U. A new copper cut-off value for diagnosis of Wilson disease? *Nat Rev Gastroenterol Hepatol*. 2015;12(9):493-494.
6. Hedera P. Wilson's disease: a master of disguise. *Parkinsonism Relat Disord*. 2019;59:140-145.
7. Leggio L, Addolorato G, Abenavoli L, Gasbarrini G. Wilson's disease: clinical, genetic and pharmacological findings. *Int J Immunopathol Pharmacol*. 2005;18(1):7-14.
8. Mak CM, Lam C-W. Diagnosis of Wilson's disease: a comprehensive review. *Crit Rev Clin Lab Sci*. 2008;45(3):263-290.
9. Lucena-Valera A, Perez-Palacios D, Muñoz-Hernandez R, Romero-Gómez M, Ampuero J. Wilson's disease: revisiting an old friend. *World J Hepatol*. 2021;13(6):634-649.
10. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2020;2021:n71. doi:10.1136/bmj.n71. Published online.

11. Nicastro E, Ranucci G, Vajro P, Vegnente A, Iorio R. Re-evaluation of the diagnostic criteria for Wilson disease in children with mild liver disease. *Hepatology*. 2010;52(6):1948-1956.
12. Sezer OB, Perk P, Hoşnut FÖ, Köse SK, Özçay F. Is it necessary to re-evaluate diagnostic criteria for Wilson disease in children? *Turk J Gastroenterol*. 2014;25:690-695.
13. Xu R, Jiang Y, Zhang Y, Yang X. The optimal threshold of serum ceruloplasmin in the diagnosis of Wilson's disease: a large hospital-based study. *PLoS One*. 2018;13(1):e0190887.
14. Merle U, Eisenbach C, Weiss KH, Tuma S, Stremmel W. Serum ceruloplasmin oxidase activity is a sensitive and highly specific diagnostic marker for Wilson's disease. *J Hepatol*. 2009;51(5):925-930.
15. Mak CM, Lam C-W, Tam S. Diagnostic accuracy of serum ceruloplasmin in Wilson disease: determination of sensitivity and specificity by ROC curve analysis among ATP7B-genotyped subjects. *Clin Chem*. 2008;54(8):1356-1362.
16. Lu Y, Liu X-Q, Wang X-H, Wang J-S. The reassessment of the diagnostic value of 24-hour urinary copper excretion in children with Wilson's disease. *Zhonghua gan zang bing za zhi= Zhonghua ganzang-bing zazhi= Chinese. J Hepatol*. 2010;18(1):49-53.
17. Ferenci P, Steindl-Munda P, Vogel W, et al. Diagnostic value of quantitative hepatic copper determination in patients with Wilson's Disease. *Clin Gastroenterol Hepatol*. 2005;3(8):811-818.
18. Yang X, Tang X, Zhang Y, et al. Prospective evaluation of the diagnostic accuracy of hepatic copper content, as determined using the entire core of a liver biopsy sample. *Hepatology*. 2015;62(6):1731-1741.
19. Cauza E, Maier-Dobersberger T, Polli C, Kaserer K, Kramer L, Ferenci P. Screening for Wilson's disease in patients with liver diseases by serum ceruloplasmin. *J Hepatol*. 1997;27(2):358-362.
20. Ceruloplasmin (Blood) - Health Encyclopedia - University of Rochester Medical Center. [https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=167&contentid=ceruloplasmin\\_blood](https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=167&contentid=ceruloplasmin_blood). Accessed October 16, 2021.
21. Zarina A, Tolmane I, Krumina Z, Tutane AI, Gailite L. Association of variants in the CP, ATOX1 and COMMD1 genes with Wilson disease symptoms in Latvia. *Balk J Med Genet BJMG*. 2019;22(2):37.

**How to cite this article:** Salman HM, Amin M, Syed J, et al. Biochemical testing for the diagnosis of Wilson's disease: A systematic review. *J Clin Lab Anal*. 2022;36:e24191. doi:[10.1002/jcla.24191](https://doi.org/10.1002/jcla.24191)