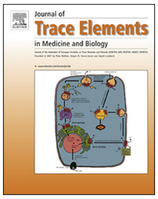
Journal of Trace Elements in Medicine and Biology 52 (2019) 288–292

Contents lists available at ScienceDirect

Journal of Trace Elements in Medicine and Biology

journal homepage: www.elsevier.com/locate/jtemb



Biochemistry

Is the measurement of copper and iron in liver biopsies reliable? Results from a pilot external quality assurance scheme[☆]

Andrew Duncan[a,*], Christopher F. Harrington[b], Anthony Catchpole[a], Andrew Taylor[b]

[a] Scottish Trace Element and Micronutrient Diagnostic and Research Laboratory, Glasgow Royal Infirmary, Castle St, Glasgow, G4 0SF, UK
[b] SAS Trace Element Centre, Surrey Research Park, 15 Frederick Sanger Road, Guildford, Surrey, GU2 7YD, UK

ARTICLE INFO

Keywords:
External quality assurance
Liver copper
Liver iron
Proficiency testing

ABSTRACT

The determination of copper and iron in liver biopsies have important roles in the diagnosis of Wilson's Disease and haemochromatosis respectively. An external quality assurance scheme is essential for analytical validation of results, however, none was available for these analytes at the time of this investigation. Accordingly a pilot scheme was established. The results of this scheme and of a corresponding questionnaire are recorded.

Twenty-nine identical sets of lyophilised certified reference materials or dried aliquots of livers purchased from local stores were distributed to 26 specialist trace element laboratories offering this clinical service. Using results returned, target values were assigned and analytical imprecision and accuracy were assessed. Laboratories were also asked to complete a questionnaire regarding details of sample preparation, analytical method and interpretation.

Accuracy was worse than ± 50% at least one result in 38% of laboratories measuring copper and 57% measuring iron. Within-batch imprecisions poorer than ± 50% were found in 20% and 23% of liver copper and iron results respectively. Accuracy was found to be statistically poorer if sample weights less than 3 mg were measured. Reference ranges were frequently absent from reports or disagreed with international guidelines.

A significant proportion of laboratories are unable to provide consistently reliable analytical performance to the extent that misdiagnosis may occur. The current diagnostic service is further compromised by the lack, or unreliability, of appropriate reference ranges provided on reports making interpretation more difficult. Without improvement, a review of current guidelines may be necessary.

1. Introduction

The determination of liver copper has an important diagnostic role in the investigation of Wilson's Disease [1,2] and it is crucial that a reliable analytical method is available [3,4]. I spite of the central role these tests may play in assisting the diagnosis, and regularly document accuracy. For this reason a pilot EQA scheme was established for liver copper.

2. Materials and methods

2.1. Participants

Thirteen laboratories known to routinely measure concentrations of copper and iron in liver were invited and eleven agreed to participate. After the first distribution of samples a further ten laboratories were recruited and for the final two distributions another three laboratories were enrolled. Over the two-year period of the scheme, twenty-six trace element laboratories in seven countries participated.

Abbreviations: EQA, external quality assurance; AAS, atomic absorption spectrometry; ICP-MS, inductively coupled plasma mass spectrometry; ICP-OES, inductively coupled plasma optical emission spectrometry; NAA, neutron activation analysis

[☆] Laboratory where work was carried out: Scottish Trace Element and Micronutrient Diagnostic and Research Laboratory, Glasgow Royal Infirmary, Castle St, Glasgow, G4 0SF, UK.

[*] Corresponding author. Current Address at: Division of Health Sciences, Royal Infirmary of Edinburgh, Little France, Edinburgh, EH16 4SA, UK.

E-mail addresses: v1adunc9@exseed.ed.ac.uk (A. Duncan), chris.harrington1@nhs.net (C.F. Harrington), Anthony.Catchpole@ggc.scot.nhs.uk (A. Catchpole), m220501@aol.com (A. Taylor).

https://doi.org/10.1016/j.jtemb.2019.01.009
Received 8 November 2018; Received in revised form 9 January 2019; Accepted 15 January 2019
0946-672X/

Table 1
Assigned values for copper and iron and the number of separate distributions circulated.

Sample	Assigned Value for Copper in µg/g (SD)	Assigned Value for Iron in µg/g (SD)	Number of Times Distributed
Muscle H4	4.3 (1.5)	51 (14)	2
Blood A13	4.8 (2.0)	2181 (530)	1
Chicken Liver	14.6 (2.8)	263 (67)	2
Bovine Kidney BCR 186	31 (5.9)	296 (55)	7
Pig Liver	31 (8.2)	1048 (194)	1
Oyster Tissue 1566a	56 (13)	328 (83)	1
Bovine Liver 1577a	159 (15)	207 (16)	1
Bovine Liver BCR 185	174 (23)	210 (57)	7
Lamb Liver 1	225 (42)	309 (52)	2
Cow Liver	342 (45)	408 (56)	2
Lamb Liver 2	562 (98)	298 (91)	2
Sheep Liver	588 (63)	317 (101)	2

2.2. Samples

The following lyophilised certified reference materials were purchased: H-4 animal muscle, A-13 animal blood, H-8 horse kidney, (International Atomic Energy Agency, PO Box 100, 1400 Vienna, Austria), 1577a bovine liver, 1566b oyster tissue, (National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, Maryland 20899, USA), BCR 185 bovine liver and BCR 186 pig kidney (European Community Bureau of Reference, Retieseweg 111, Geel, Belgium). Fresh samples of chicken, sheep, cow, pig and lamb liver were purchased from local stores. Aliquots of these were oven-dried at 37 °C until a constant weight was reached and placed in small sealed tubes for distribution. Ten additional samples, taken from different sites of each liver, were analysed prior to distribution to determine the homogeneity of the fresh liver samples. In each case copper and iron concentrations were reproducible with a coefficient of variation of less than 5%.

2.3. Design

Over a two-year period, 29 test samples in 8 distributions (3 samples in the first 3 distributions and 4 samples in the last 5 distributions) were posted to the participants. On 10 occasions, pairs of the same sample were sent in the same distribution to investigate within-batch imprecision and on 7 occasions, pairs of the same sample were sent in different distributions to investigate between-batch imprecision.

The weights of samples distributed (mean: 6.6 mg, range: 0.6–26 mg) were greater than would typically be received in clinical laboratories; for example, over the past year the mean weight of the 31 liver biopsies received for analysis in Scotland was 2.7 mg (range: 0.3–9.1 mg) and in the Guildford laboratory over the course of the last seven years, of the 75 samples received the mean weight was 2.5 mg (range: 0.2–8.8 mg). Smaller amounts of sample were distributed in 7 of the 29 sets of samples distributed in order to assess the analytical accuracy of low-weight samples. Laboratories were asked to prepare and analyse the samples as would be done routinely, calculate results as µg/g of dried weight of tissue, and return the results and weights of samples analysed (in mg) within three months.

In order to obtain assigned values, results which were greater than double or less than half of the median were considered to be blunders and excluded; results outwith ± two standard deviations were then excluded and the mean of the remaining results was taken as the assigned value.

To assess the accuracy, each result was scored from 1 to 5 based on the percentage bias from the assigned value: 5 (± 0 to ± 5%), 4 (> ± 5% to ± 10%), 3 (> ± 10% to ± 20%), 2 (> ± 20% to ± 50%) and 1 (> ± 50%). In order to calculate inaccuracy of individual laboratories, blunders were initially excluded

To assess if there was a difference in the accuracy of lyophilised samples compared with dried liver samples, the copper results were compared by student's *T*-test. Samples with low concentrations tend to have greater inaccuracy and so percentage biases for each sample type were initially calculated for each sample type; it was found that the average bias decreased as copper concentrations increased to a plateau of 100 µg/g so only samples above this concentration were used for this comparison. Copper rather than iron results were used because of the superior analytical performance of its measurement.

To assess imprecision, the percentage difference between pairs of results divided by the mean, was calculated. Imprecision was scored from 1 to 5 in the same way as accuracy.

Questionnaires were sent to participants requesting details of sample preparation, if a minimum weight of liver is required for routine analyses and, if so, the weight below which analysis is not carried out, analytical method, reference ranges quoted on reports, diagnostic cut-off concentrations quoted on reports, and if an interpretation of results was given.

3. Results

Of the 26 laboratories one laboratory did not report iron results. Seven of the laboratories dropped out over the course of the scheme. Results were not returned from 12 of 118 (10%) sets of samples sent.

The assigned value of each of the samples is shown in Table 1.

3.1. Accuracy

Table 2 shows the accuracy of results based on the scoring criteria described.

There was no significant difference in accuracy of copper measurement between lyophilised samples and liver samples. The mean results were 8% (SD: 9%) for 129 lyophilised samples and 12% (SD: 12%) for 109 liver samples.

The dry weights of 292 samples were reported of which 70 were less than or equal to 3 mg. The accuracy of samples with weights less than equal to 3 mg was found to be poorer than those with weights over 3 mg ($p = 0.02$); average bias from the assigned value of 33.3% vs 15.4% respectively.

Table 2
Accuracy of copper and iron measured as the percentage bias from the assigned value.

Score	Accuracy	Copper (n = 436)	Iron (n = 415)
5	± 0 to ± 5%	124 (28%)	76 (20%)
4	> ± 5% to ± 10%	106 (24%)	88 (21%)
3	> ± 10% to ± 20%	112 (26%)	108 (32%)
2	> ± 20% to ± 50%	35 (8%)	79 (14%)
1	> ± 50%	59 (21%)	64 (13%)

Table 3
Imprecision of copper and iron analysis expressed as percentage deviation.

Score	Interpretation	Within-batch imprecision		Between-batch imprecision	
		Copper (n = 156)	Iron (n = 148)	Copper (n = 81)	Iron (n = 78)
5	0 to 5%	80 (51%)	47 (32%)	25 (31%)	15 (19%)
4	> 5% to 10%	34 (22%)	34 (23%)	13 (16%)	15 (19%)
3	> 10% to 20%	29 (22%)	33 (29%)	20 (25%)	19 (24%)
2	> 20% to 50%	11 (4%)	27 (11%)	13 (16%)	22 (28%)
1	> 50%	2 (1%)	7 (5%)	9 (11%)	7 (9%)

3.2. *Imprecision*

Within-batch sample pairs: A total of 156 pairs of results was obtained on replicate samples sent in the same distribution for copper and 148 paired results for iron. The within-batch percentage difference between pairs of results was scored at 4 or 5 (≤ 10%) in 73% of the samples for copper and 55% for iron. Scores of 1 or 2 (> 20%) were found in 5% of copper results and 16% for iron results. See Table 3.

A total of 81 pairs of copper results and 78 pairs of iron results was obtained from duplicate samples between batches. Results of three pairs (8%) were excluded because of possible blunders. The percentage difference for between-batch imprecision was scored at 4 or 5 (≤ 10%) for copper and iron in 47% and 38%, respectively. Imprecision was scored as 1 or 2 (> 20%) for between-batch analyses in 27% and 37% for copper and iron.

3.3. *Analytical technique*

Details of the analytical technique used were received from 21 laboratories. Performance in terms of accuracy (Table 4) and imprecision (Table 5) appears to be superior for atomic absorption spectrometry (AAS) and inductively coupled plasma mass spectrometry (ICP-MS) compared with inductively coupled plasma optical emission spectrometry (ICP-OES) and neutron activation analysis (NAA) although the number of laboratories was insufficient to perform meaningful statistical analysis.

There was a considerable systematic bias of greater than ± 20% in one lab that measured copper (mean of -37%) and four laboratories that measured iron (means of -23%, -37%, -48% and 48%).

3.4. *Blunders*

15 of 448 (3.3%) copper results and 26 of 412 (6.3%) iron results were identified as blunders presumably as a result of analytical, chemical or clerical errors.

Of the 26 laboratories which measured copper and 25 which measured iron, blunders were present in 9 (35%) and 19 (76%) respectively.

Table 4
Accuracy of copper and iron analysis according to method.

Method	Number of laboratories	Number of samples		Mean percentage bias from assigned value	
		Copper	Iron	Copper	Iron
AAS	3	38	36	12%	17%
Flame AAS	3	25	10	12%	34%
Graphite AAS	3	44	26	13%	22%
ICP-MS	7	123	149	11%	14%
ICP-OES	4	61	69	17%	17%
NAA	1	12	14	26%	29%

3.5. *Questionnaire responses*

Responses were received from 18 of 28 (64%) laboratories although many of them were incompletely answered.

Upper reference limits for liver copper (number of laboratories) were reported in 12: 25 μg/g (1), 32 μg/g (1), 40 μg/g (1), 50 μg/g (7), 57 μg/g (1) and 160 μg/g (1). US and European Wilson's Disease guidelines both quote an upper reference limit of 50 μg/g.[1,2] Cut-off limits for liver copper (number of laboratories) suggestive of Wilson's Disease were reported in 5 (19%): 250 μg/g (3), 254 μg/g (2). These figures are in accordance with US and European guidelines respectively for Wilson's Disease [1,2].

Upper reference limits for liver iron (number of laboratories) were reported as 300 μg/g (1), 1400 μg/g (2), 1500 μg/g (4), 1700 μg/g (1), 1787 μg/g (1), 2000 μg/g (1), 2010 μg/g (2). The figure quoted in the US guidelines is 1500 μg/g. Only one laboratory quoted a cut-off figure (≥ 3351 μg/g) suggestive of haemochromatosis. This compares with US guideline figures of > 2000 μg/g and > 10,000 μg/g for asymptomatic and symptomatic haemochromatosis respectively. Only two laboratories quoted an hepatic iron index (1.9 and 2.0 μmol/g/age) as indicative of haemochromatosis. US haemochromatosis guidelines quote a cut-off of > 1.9 μmol/g/age.

The numbers and types of analytical techniques used are shown in Table 4.

All laboratories used different digestion methods. These varied by the type, purity, volume and concentration of digestion agent used and the time, temperature and type of heating that the samples were subjected to. The types of digestion agent used (numbers of laboratories) were nitric acid (12), nitric acid/perchloric acid (3), nitric acid/sulphuric acid (1), tetramethyl ammonium hydroxide (1) and trypsin (1): purities were SupraPur (5), UltraPur (2), NormaPur (2), Plasma Pure (1) and GPR (1): volumes added for digestion varied from 200 μL to 5 mL, commonest 500 μL (8); concentrations varied from 50% to concentrated, commonest 65% nitric acid (7); times of heating varied from 3 min to 24 h, commonest overnight (4); temperatures of heating varied from room temperature to 400 °C, commonest being 60 °C (4). Heating methods used were heating block (8), oven (2), microwave oven (2), sand block (2), sonication (1) and water bath (1).

The lowest acceptable weights of biopsy for analysis (number of laboratories) were quoted in 8 (28%): 0.1 mg (1), 0.2 mg (1), 1 mg (1), 5 mg (3), 10 mg (1).

4. Discussion and conclusion

There are several factors which may make the analysis of trace elements difficult. By definition, the concentrations of trace elements in biological samples are very low, typically less than μg/g or μg/mL; they may be ubiquitous in the environment and so analysis may be compromised by contamination during sample collection, preparation and analysis; the amount of sample available for analysis may be very small, limiting the number of replicate measurements that can be made; techniques to determine low concentrations are complex and require experienced, trained analysts to maintain optimum performance; reliable trace element results from biological samples are best achieved in laboratories that regularly analyse specimens and that have established the necessary expertise to accommodate these challenges [5,6]; external quality assessment (proficiency testing) schemes (EQAS) are available to validate methods and monitor on-going performance but only for clinical fluids such as serum and urine; in addition, the only tissue certified reference materials available are lyophilised powders which do not match the matrix of real specimens.

At the time of this study, no EQA scheme for liver copper and iron was available and so a pilot scheme was established in order to assess the analytical quality of these measurements. The results of this pilot scheme indicate that several clinical laboratories are unable to report accurate results and in 38% of laboratories measuring copper and 57%

Table 5
Imprecision of copper and iron analysis according to method.

Method	Number of labs	Number of samples	Mean (SD)	Number of labs	Number of samples	Mean (SD)
	Within Batch Imprecision			Between Batch Imprecision		
AAS	2	10	1.9% (1.7)	2	7	1.1% (0.8)
Flame AAS	1	8	3.1% (4.2)	1	3	0.6% (0.5)
Graphite AAS	3	14	1.5% (1.2)	3	6	3.7% (3.5)
ICP-MS	6	43	1.7% (1.7)	6	29	2.4% (1.9)
ICP-OES	3	19	3.0% (4.2)	4	12	5.3% (6.4)

Table 6
Recommendations for optimum clinical service for liver copper and iron.

Establish an internal quality control scheme covering high and low concentrations (see Table 1) using locally sourced liver samples.

Report weight of sample with a suitable comment noting the increased possibility of analytical inaccuracy if it is below 3 mg (eg. Please note results are increasingly inaccurate as biopsies decrease in weight below 3 mg).

Quote, in reports, the upper reference limits defined in US and European guidelines i.e. liver copper: 50 µg/g [1,2]; liver iron 1500 µg/g [3].

Quote, in reports, both of the lower liver copper diagnostic limits for Wilson's Disease as defined in European guidelines [1].

- 250 µg/g (83% sensitive, 99% specific for Wilson's Disease).
- 75 µg/g (97% sensitive, 95% specific for Wilson's Disease).

Quote, in reports, the lower hepatic iron index for haemochromatosis as defined in US guidelines [3].

- 1.9 µmol/g/age (consistent with asymptomatic hereditary haemochromatosis).

If good results from EQA scheme cannot be attained consider referring samples to specialist laboratory with acceptable analytical performance.

Report results electronically to reduce the risk of clerical blunders.

measuring iron, accuracy was worse than ± 50% for at least one result. Similarly, poor between-batch imprecisions, of worse than 50%, were found in 20% and 23% of liver copper and liver iron results respectively.

It is important that the weight of biopsy can be reliably measured since copper and iron results are reported with respect to the dry weight of tissue. We have shown that the accuracy of results is decreased by an average of 15% if sample weights are less than 3 mg and that the degree of inaccuracy increases as the weight of biopsy is reduced further. Consequently the inaccuracy of results reported in routine practice will almost certainly be greater because of the smaller sample weights compared to those used for this study (2.6 mg vs 6.6 mg).

Only eight laboratories had a policy regarding the lowest acceptable weight for analysis. Two laboratories suggested unrealistically low cut-offs of 0.1 mg and 0.2 mg and four laboratories suggested cut-offs (5 mg and 10 mg) that are higher than the weights that would usually be received. The questionnaire sent did not ask if laboratories also routinely report the dry weight of biopsy, however, given the effect of low sample weights on accuracy, it would seem sensible to do so and also indicate in a standard comment, as one laboratory did, that analytical inaccuracy increases with decreasing biopsy weights below 3 mg.

For clinicians to correctly interpret results, laboratories should cite well-defined upper reference limits, lower reference limits consistent with Wilson's Disease and haemochromatosis, and the dry weight of biopsy received. However, only 7 and 4 laboratories quote upper reference limits for copper and iron respectively that accord with European and US guidelines [1,2] (liver copper of 50 µg/g and liver iron of 1500 µg/g). One laboratory quotes an inappropriately high liver copper reference limit of 160 µg/g. Similarly, only a minority of laboratories quotes lower cut-off concentrations that are consistent with Wilson's Disease and haemochromatosis (19% and 33% respectively). No laboratory quoted the value of 1.2 µmol/g (equivalent to 75 µg/g) which is given in the European guidelines for Wilson's Disease [1]. This concentration improves the diagnostic sensitivity (from 83.3%–96.5%) at the expense of a small decrease in specificity (from 98.6%–95.4%) [7]. Hepatic iron concentrations increase with age [8,9] and so it is useful to correct for this by calculating the hepatic iron index, the iron concentration (in µg/mol) divided by the age in years, however, only 2 (8%) laboratories currently do so.

While AAS and ICP-MS methods appear to give more accurate and precise results than do those using ICP-OES and NAA, no definitive conclusions can be made because of the small numbers of participating laboratories.

The request rate for liver copper analysis is very low (less than weekly in Scotland and less than monthly in Guildford) and so it is likely that results are more often calculated and reported manually rather than electronically as is the case with most other clinical biochemistry analytes. As a result, the incidence of blunders is likely to be disproportionally higher. In this study, the blunder rates were 3.3% and 6.3% for liver copper and liver iron respectively which may result in the infrequent but important risk of diagnostic misclassification. For example, five copper results from four laboratories were sufficiently inaccurate to result in misdiagnosis; i.e. results within the reference range (< 50 µg/g) were reported for three samples with assigned values consistent with Wilson's Disease (> 250 µg/g) while results of 554 and 573 µg/g which are highly suggestive of Wilson's Disease were reported for two samples whose assigned values were within the reference range.

The experience of this pilot scheme shows that even in specialist laboratories offering an analytical service for measuring concentrations of copper and iron in liver, the accuracy and between-batch imprecision of results is often inadequate. The weights of most samples distributed in this pilot scheme were higher than would usually be received and given the finding that low sample weights increase inaccuracy, in routine practice the robustness of analysis is likely to be even poorer than found in this study. Since the amount of liver biopsy received precludes the opportunity for a repeat analysis, it is imperative that an EQA scheme is available and that trace element laboratories participate to improve the quality of their measurements. As a consequence of the results from this study an international EQA scheme was established (Solid matrix scheme, UK NEQAS for Trace Elements, phone: 0044-1483 689022, rsc-tr.guildford-eqa@nhs.net). We recommend that laboratories which measure liver copper and iron enrol in this scheme and adopt the following recommendations (Table 6) to improve and maintain the quality of their results. If the analytical performance in this new scheme does not improve significantly, it may be necessary to review the diagnostic role of liver copper and iron in current guidelines.

Declarations of interest

None.

Funding

None.

References

[1] B.R. Bacon, P.C. Adams, K.V. Kowdley, L.W. Powell, A.S. Tavill, Clinical practice guidelines: Wilson's disease. European association for the study of the liver, Hepatology 54 (2011) 328–343.

[2] E.A. Roberts, M.L. Schilsky, American Association for the Study of Liver Disease, Practice guidelines. Diagnosis and treatment of Wilson Disease: an update,

Hepatology 47 (2008) 2089–2111 Eve A, Michael L.

[3] B.R. Bacon, R.S. Britton, Hemochromatosis and other iron storage disorders, in: E.R. Schiff, M.F. Sorrell, W.C. Maddrey (Eds.), Diseases of the Liver, 9th ed., Lippincott Williams & Wilkins, Philadelphia, 2003, pp. 1195–2015.

[4] B.R. Bacon, P.C. Adams, K.V. Kowdley, L.W. Powell, A.S. Tavill, Diagnosis and management of hemochromatosis: practice guideline by the American Association for the Study of Liver Diseases, Hepatology 54 (2011) 328–343.

[5] CLSI, Control of Preanalytical Variation in Trace Element Determinations; Approved Guidelines, second edition, (2001) CLSI document C38-A2. (ISBN 1-56238-000). CLSI, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898, USA.

[6] A. Berlin, G. Mattiello, A. Taylor, Quality assurance for aluminium analysis, Trace Elements in Human Health and Disease. WHO Environmental Health Report 20, WHO, Copenhagen, 1987, pp. 124–127.

[7] P. Ferenci, P. Steindl-Munda, W. Vogel, W. Jessner, M. Gschwantler, et al., Diagnostic value of quantitative hepatic copper determination in patients with Wilson disease, Clin. Gastroenterol. Hepatol. 3 (2005) 811–818.

[8] M.L. Bassett, J.W. Halliday, L.W. Powell, Value of hepatic iron measurements in early hemochromatosis and determination of the critical iron level associated with fibrosis, Hepatology 6 (1986) 24–29.

[9] P.C. Adams, C. Bradley, A.R. Henderson, Evaluation of the hepatic iron index as a diagnostic criterion for genetic hemochromatosis, J. Lab. Clin. Med. 130 (1997) 509–514.