Acta Ophthalmologica

Colour vision classification – comparing CAD and CIE 143:2001 International recommendations for colour vision requirements in transport

Vilhelm F. Koefoed,[1] Tyler Miles,[2] John B. Cason[2] and Ray Troche[2]

[1]Norwegian Centre for Maritime and Diving Medicine, Norwegian Armed Forces, Joint Medical Services, Department of Global Public Health and Primary Care, University of Bergen Faculty of Medicine and Dentistry, Bergen, Norway
[2]Naval Medical Center San Diego, Navy Refractive Surgery Center, San Diego, CA, USA

ABSTRACT.
Purpose: **To evaluate the colour vision severity classification standard 'CIE 143:2001 International recommendations for colour vision requirements in transport' (CIE 143:2001), which has become out of date because of the lack of commercial availability of required colour vision tests.**
Methods: **One-hundred-five subjects had colour vision tested and colour vision severity classified according to a modified CIE 143:2001 algorithm that included pseudoisochromatic plates (Ishihara's test and Hardy Rand Rittler (HRR) 4th edition), Optec 900 lantern and Farnsworth D-15. Subject's results and colour vision severity classification were compared to performance and colour vision severity classification on the computerized 'Colour Assessment and Diagnosis' (CAD) test.**
Results: **According to CIE 143:2001, using Ishihara's test, Optec lantern and Farnsworth D 15, 11 subjects (10%) were category I (normal), 16 (15%) were category II (mild), 48 (46%) were category III (poor), and 30 (29%) were category IV (severe). Classified by CAD score, 10 (10%) were category I, 11 (10%) were category II, 41 (39%) were category III, and 43 (41%) were category IV. The correlation between the two estimates of the severity of colour vision loss (i.e. CIE 143:2001 and CAD) was high, with a Kendall's Tau test of 0.81 ($\tau = 0.81$ $p < 0.001$). A suggested CIE 143:2001 classification including new CAD score limits improves the classification correlation to 0.90 ($\tau = 0.90$ $p < 0.001$) for all diagnoses.**
Conclusion: **The colour vision severity classification standard 'CIE 143:2001 International recommendations for colour vision requirements in transport', has not implemented new diagnostic tools with better accuracy. We propose three possible revisions to the CIE 143:2001 algorithm, based on the availability of CAD: (1) Replacing the current CIE 143:2001 algorithm using new CAD threshold limits, (2) Use of CAD as a secondary test to Ishihara's test and HRR or (3) Revising the current CIE 143:2001 algorithm using Ishihara's test, HRR, Optec 900 and FD15.**

Key words: colour vision deficiency – colour vision classification – severity – CAD – CIE 143:2001

Acta Ophthalmol.

doi: 10.1111/aos.14450

Introduction

Historical background

Colour vision testing standards received attention in 1855 when Dr George Wilson (1855) pointed out the necessity of using coloured signals secondary to non-coloured signals at railway stations and aboard ships. However, it was not until 1877 that the first regulations were adopted in the European railroad industry. After a railway accident in Sweden in 1875, the Swedish ophthalmologist Dr Frithiof Holmgren (1878) introduced the theory of colour vision deficiency (CVD) as a significant cause of the incident. However, this explanation has since been disputed (Mollon and Cavonius 2012). This accident was, nevertheless, one of several contributing factors for the introduction of colour vision standards in European railroad and maritime industries. The increased awareness of CVD and the increased use of coloured signals created public demand for better safety in public transportation (Vingrys & Cole 1986; French & al. 2008).

From 1877 and onward, the railway and maritime transportation industries adopted colour vision testing using varying standards and methods. In 1919, the International Civil Air Navigation Authority (ICAN) set CV standards for the aviation industry, and in the mid-1930s, CV standards were set for road transportation in Britain (Vingrys & Cole 1986; French & al. 2008).

Except for the British road transportation CV standard, which was ended in 1960; few questions have been made for the validity of CV standards in other occupations. Joint Aviation Authorities aviation regulations have adopted the findings of a task performance study by the Safety Regulation Group (2009), allowing subgroups of deutan and protan CVD to be commercial pilots. In 2001, the International Commission on Illumination (CIE) recommended a new standard, 'CIE 143:2001 International recommendations for colour vision requirements in transport' (CIE 143:2001) (CIE 2001), for CV testing and classification of the result by classifying the CV in four groups (CIE 1-3 and non-classifiable (CIE 4)) (Table 1). This recommendation is now considered outdated according to Bailey & Carter (2016), and Carter & Barbur (2015) have recommended changes in the test protocol for personnel on lookout duties.

CIE 143:2001 colour vision classification

Colour vision classification, according to CIE 143:2001 (Fig. 1), includes several different tests, including pseudoisochromatic plates (PIP), spectral anomaloscopy, Medmont, lanterns and Farnsworth D15 (FD15). The most commonly used PIP test is Ishihara's test, which employs a range of designs, such as transformation, vanishing or hidden digits. The Ishihara's test employs luminance and yellow–blue (YB) colour noise and is therefore limited to red–green (RG) deficiency. An alternative PIP test is the Hardy Rand Rittler (HRR) 4th edition. This test employs vanishing type plates with varying difficulty for CVD subjects and includes YB CV sensitivity. Hardy Rand Rittler (HRR) makes it possible to diagnose and quantify the CV. Spectral anomaloscopy is based on colour matching and is accepted as the diagnostic reference test for the diagnosis of RG CVD. The Medmont test was designed to identify and quantify CVD, thus differentiating between deutans and protans. Protans have reduced visual sensitivity for red light. The Medmont test is no longer commercially available. Lanterns simulate coloured light signals used in the transportation industry. CIE 143:2001 administers Holmes Wright type B (HW-B) lanterns as a secondary test in addition to Ishihara's test for identifying CVN subjects (Fig. 1). Several other lanterns, like Holmes Wright type A (HW-A), OPTEC 900 and Farnsworth, are used for other CIE 143:2001 classifications. The HW-B lantern was explicitly developed to illustrate the requirements for CV for navigation duties in the maritime transport industries as described by IALA (2017). Holmes Wright type B (HW-B) is an improvement of the Board of Trade's Lantern of 1913, the first approved maritime lantern for testing CVD. Lanterns can be classified according to the chromaticity chart with relevance to the CIE 1931 standard colourimetric observer (CIE 2019). Holmes Wright type A and B are Green A lanterns while others are Green B lanterns, and the results cannot be used interchangeably. The last test used in CIE 143:2001 is the dichotomous FD15, a hue-test. It is comprised of 16 discs of different coloured surfaces, which must be arranged in ascending order of the hue. CVN subjects and mild CVD subjects can organize the colours correctly, while dichromat CVDs will demonstrate the lack of ability for distinguishing coloured surfaces. The pattern of errors indicates the diagnosis of CVD (Dain SA & al. 2019)

CIE 143:2001 uses the tests described above to classify CV into four mutually exclusive groups (CIE 1-3 and non-classifiable (CIE 4)) (Table 1, Fig. 1).

CIE colour vision standard 1 (CIE 1) is considered normal colour vision and is regarded as the appropriate standard when coloured signals must be recognized at long distances or under adverse visibility conditions. It also applies when more than three surface colours are used as a primary and essential code of conveying information. The fail criteria for CIE 1 are errors on three or more plates on the Ishihara's 38 plates edition. If the result of the Ishihara's test is not conclusive a secondary test, another PIP, anomaloscopy or HW-B, may be performed. CIE colour vision standard 2 (CIE 2) passes subjects with mild CVD, but who have demonstrated the ability to identify coloured signals correctly. A failed Medmont test or lantern test, other than HW-B, are exclusion criteria

Table 1. This table describes the CIE 143:2001 colour vision standards that are used for classification and relate to the International recommendations for colour vision requirements in transport.

CIE colour vision standard 1
This standard requires normal colour vision.
It is the appropriate standard when:
Coloured signal must be recognised at long distance or under adverse visibility conditions, or surface colour codes with more than three colours are used as a primary and important means or conveying information on computer screens or other visual information displays
And failure to see or identify a coloured signal or colour code is likely to cause an operational error or accident,
High social, economic or environmental costs likely to be associated with accidents or there is a community expectation of the highest standards of safety.
CIE colour vision standard 2 (defective colour vision A)
This standard passes persons who have a mild colour vision deficiency but have demonstrated ability to identify coloured signals correctly. It fails those persons with a protan defect colour vision who have a very reduced ability to see red signal lights.
It is the appropriate standard when recognition of coloured lights or other coloured codes is important to safe operation, or red signals has to be seen.
And significant social, economic or environmental costs may be associated with accidents, or there is a community expectation of high standards of safety.
CIE colour vision standard 3 (defective colour vision B)
This standard passes persons who have a colour vision deficiency but have demonstrated ability to recognise colour codes at short distances.
It is the appropriate standard when:
Surface colour codes, including colour codes on a computer screen displays and large signal lights on control panels, have to be recognized at short distances under good conditions of visibility, And
Economic or environmental costs may be associated with operational errors or accidents
CIE non-classified colour vision
This standard is appropriate when the persons have demonstrated a lack of ability to recognize coloured signals or codes.

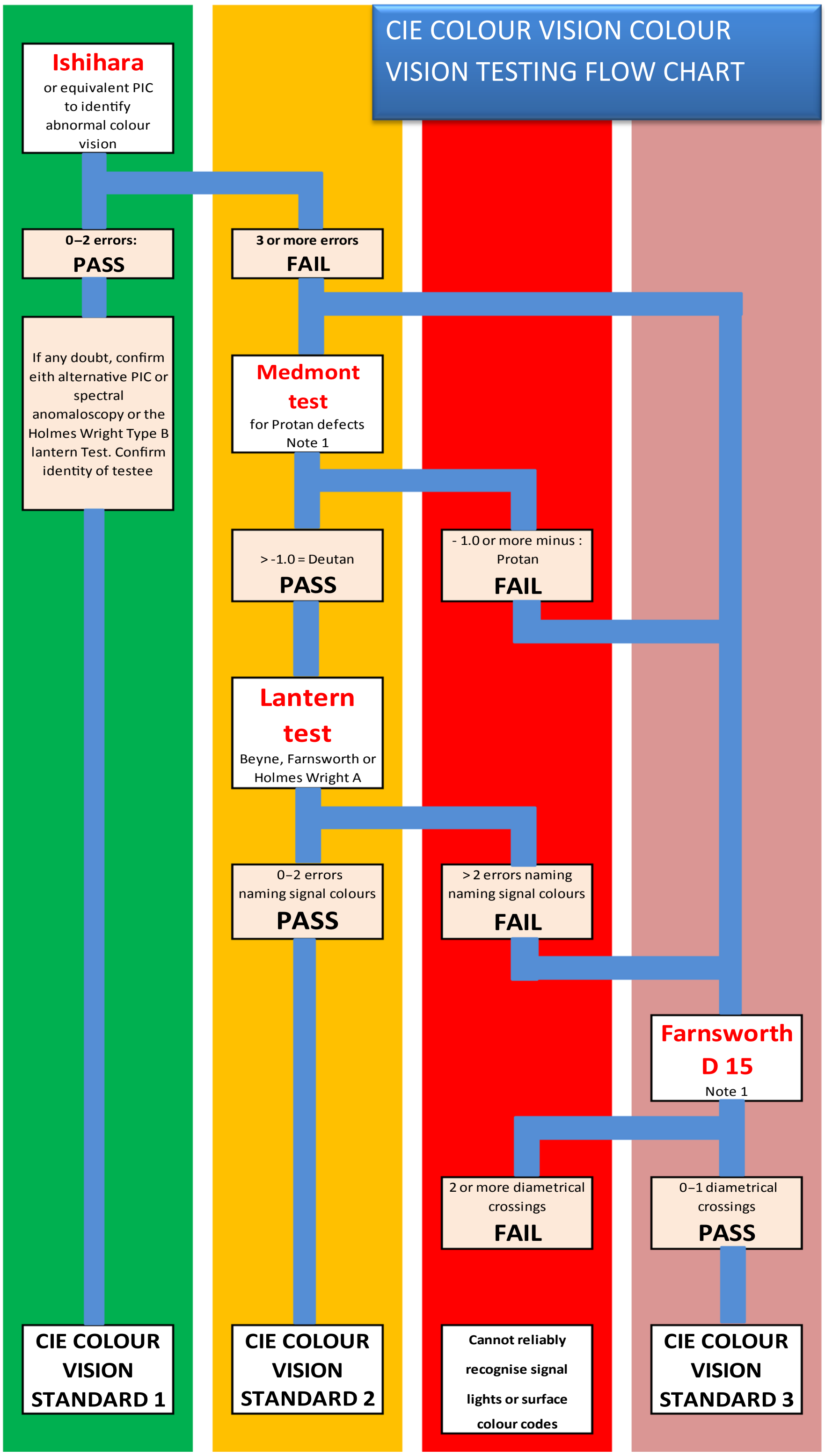

Fig. 1. Recommended flow chart for CIE 143:2001 colour vision classification (CIE 143:2001).

for CIE 2. The protocol allows for 1–2 errors, depending on the lantern used when naming the signal colour combinations. In CIE colour vision standard 3 (CIE 3), the subject demonstrates the ability to recognize surface codes (e.g. coloured surfaces) at short distances. The requirement is to pass FD15 with less than two diametrical crossings on the score sheet. Anyone who fails CIE 3 is considered CIE non-classified standard (CIE 4) because they are unable to recognize signal lights or surface colour codes reliably.

Colour vision testing with the CAD test

In *Colour vision requirements in visually demanding occupations* Barbur and Rodriguez-Carmona (2017), introduce a new colour vision severity classification, CV 1-5. The severity classification is based on 'The Colour Assessment and Diagnosis' (CAD) test that is implemented on a calibrated visual display and consists of coloured stimuli of precise chromaticity and saturation. Colour Assessment and Diagnosis (CAD) makes it possible to establish reliable estimates of RG and YB colour thresholds. The CAD test has demonstrated close to 100% sensitivity and specificity in detecting CVD and in classifying the type of deficiency involved (Marechal & al. 2018; Barbur & Rodriguez-Carmona 2017). The severity of both RG and YB colour vision loss is quantified in CAD units as Standard Normals (SN)[1]. The severity classification consists of five categories. 'Normal' CV (CV 1) includes all subjects below the upper normal age-related threshold limit, for all practical purposes close to RG 2.25 SN (Carter & Barbur 2015). 'Functionally normal' CVD (CV 2) are applicants that exhibit almost normal RG colour discrimination (RG < 2.35 SN) and pass the HW-A with zero errors. Just over 6% of the deutan population will pass this test. The third category is 'Safe' CVD (CV 3) (RG 2.36 - 4 SN). This higher limit is sufficient to pass all CVN and about 22% of deutan subjects. This threshold matches the percentage of deutans who pass the HW-A lantern or equivalent lantern when using the CIE recommended protocol for use with this lantern. The fourth category is 'Poor' RG CV (CV 4) (RG 4.01- 12 SN). Subjects in this category will be able to make use of and cope with saturated RG colours on visual displays. However, they will typically take longer to complete colour-related tasks and may also be less accurate. The fifth classification is 'Severe' RG colour deficiency (CV 5) (RG > 12 SN). The vast majority in this group have very little use of RG colour signals and have to rely mostly on YB colour differences.

[1]One Standard Normal is based on the median RG and YB threshold values measured in 330 healthy, young subjects.

Subjects and methods

Study subjects

Subjects in the study were all clients referred for assessment of colour vision or contrast sensitivity at Haukeland university hospital, Dept. of occupational medicine. The department is recognized by the Directorate of seafarers as one of two national centres to perform secondary colour vision assessments.

Colour vision assessment in the study

Colour vision was assessed following the CIE 143:2001 protocol with some adjustments and by CAD (City Occupational Ltd). Several of the tests described in CIE 143:2001 are no longer commercially available; this includes the Medmont test and most lanterns except Optec 900. Furthermore, anomaloscopy was not included in this study. In this study, we used the 16 first plates of Ishihara's 24-plate ed. leaving plate number one out. We also used the HRR 4th edition (RichmondProducts.com). Optec 900 (stereooptical.com) and Farnsworth D15 (good-lite.com) were the two other test methods for CIE 143:2001. We used the Colour Test Daylight Illuminator with 6280-degree K light for illumination (good-lite.com) when testing Ishihara's test, HRR and FD15. The CAD test was performed in mesopic light levels in a room shaded from direct daylight. After performing the learning module, they performed the 'definitive CAD' YB test followed by RG test. Results were reported in CV diagnosis (i.e. protan, deutan or colour vision normal) and SN units. All subjects were tested with all CV tests.

Colour vision classification

The subjects were classified by CV according to each method. Colour Assessment and Diagnosis (CAD) diagnosis and quantification (SN) was used as the gold-standard reference. CIE 143:2001 classification was based on Ishihara's test, HRR, Optec 900 and FD15 tests. An Ishihara's test score with less than three errors was considered a pass. If subjects made less than two errors per nine lantern observations on Optec 900, the test was classified as a pass. FD15 was classified as a pass if there were less than two diametrical crossings on the score sheet. For HRR, no errors on plates 5–10 indicated CVN. The highest number of counts on either protan or deutan plates (plates 11–20) decided the CV diagnosis. If the numbers were equal, the diagnosis was set to non-classified. Hardy Rand Rittler (HRR) quantifies the CV in four groups: normal CV and mild, moderate and strong CVD, respectively HRR 1–4 (14). Hardy Rand Rittler (HRR) was also scored and classified by weighting (HRR-w) the score of each plate (plates 7–20) from the first red–green plate, #7 (Bailey & al. 2004). This weighted value is progressively more significant for the plates as they decrease in difficulty. The total score for protan (negative) and deutan (positive) plates was summed up, and this determined the diagnosis; a negative score indicated protan and a positive score deutan. The lowest score of either protan or deutan indicated the severity of the CVD. CAD CV 1-5 was collapsed to CV 1-4 for statistical purposes (new CV 1 includes CAD $\leq$ 2.35 SN, thus removing the original CV 2). Severity classification into CIE 1-3 and non-classifiable (CIE 4) was done according to the modification of the CIE 143:2001 protocol. The Medmont test is no longer available and was substituted by HRR to identify protan CVD.

Statistical analyses

Visual descriptors and colour vision classifications; CIE 1-4, CV 1-4 and HRR 1-4, HRR-w 1-4, were described by descriptive statistics. Kendall's tau and McNemar's tests compared colour vision classifications. Receiver operating characteristic (ROC) curves were administered for Optec 900 and FD15. The Youden index is defined as (sensitivity + specificity - 1) and is a descriptor of ROC test performance. The Youden index helps identify the optimal cut point that maximizes both specificity and sensitivity. Test accuracy will be defined by (sensitivity*prevalence) + (specificity)*(1-prevalence) (Baratloo & al. 2015). The Shapiro–Wilk test tested for normality of visual data. The level of significance was set at 0.05. Statistical analyses were performed using IBM SPSS Statistics version 25.0 (IBM corp., Armonk, NY, USA).

Research ethics

This study adhered to the Declaration of Helsinki. Subjects were informed about the objectives and conditions of the research and signed a consent form. The Regional Committee for Medical Research Ethics, Western Norway approved the study protocol, ref. 2018/657/REK vest. The test subjects were not paid for participating in the study, and they could withdraw from the study at any point.

Results

Subjects

The study included 105 subjects, six female and 99 men. The mean age was 28.7 years (range 15–69; SD 13.5). The mean age of the CVN group exceeded the CVD group by nine years ($p = 0.33$). There was no significant difference in age between the genders ($p = 0.48$) nor the CVD groups ($p = 0.57$). The largest employment group was men seeking to work in the deck department ($n = 41$) followed by machine department ($n = 23$) and in fishery ($n = 17$). A complete set of colour vision measurements were obtained from all 105 subjects (Table 2). None of the visual tests showed a normal distribution according to the Shapiro–Wilk test.

Colour vision performance

Ten subjects (9.5%), including four women, were CVN (SN 0.89–1.47; SD 0.18) by all tests, In addition to these ten subjects, Ishihara's test also

Table 2. Colour vision classification by Ishihara's test, CAD, HRR, weighted HRR (HRR-w) and FD15. In HRR-w, each plate receives a specific value, and the sum of value defines the diagnosis.

Colour vision	Ishihara	CAD	HRR	HRR-w	FD15
Normal (CVN)	11	10	10	10	74
Deutan	0	71	83	81	17
Protan	0	24	9	14	7
Non-classifiable	94	0	3	0	7
Total	105	105	105	105	105

classified one CVD as CVN (Table 2). According to CAD, 71 subjects (67 %), including one woman, were deutan (SN 2.90–30.69; SD 7.52) and 24 subjects (23 %), including one woman, were protan (SN 6.26–26.64; SD 7.26) colour vision deficient. One subject was found to be both deutan and mild tritan CVD (YB; SN 2.60), probably due to glaucoma. For this subject, only his deutan CVD was included in the results. Farnsworth D15 classified 74 subjects as normal, 17 deutan, seven protan and seven failed the test without precise classification of the CVD.

Hardy Rand Rittler (HRR) classified ten CVN, 83 deutan and nine protan CVD. Three subjects were not classified diagnostically – one deutan (3.61 SN) and two protans (both 7.03 SN). HRR-w found ten to be CVN, 81 deutan and 14 protan. The main effect of weighting the HRR plates was an increase in the number of protans by four and that all subjects were classified diagnostically.

Colour Assessment and Diagnosis (CAD) diagnosis correlated ($\tau = 0.74$, $p < 0.001$) with HRR diagnosis showing the significant discrepancy between HRR (which diagnosed nine protans) and CAD (which diagnosed 24 protans). HRR-w diagnosis correlated ($\tau = 0.79$, $p < 0.001$) with CAD. In the subgroup of CAD-diagnosed protans (24), the HRR-w classification indicated 13 protans instead of the nine in HRR.

Ninety-four (90%) subjects failed, and 11 (10%) subjects passed the Ishihara's test. Ishihara's test (pass/fail) and HRR (pass/fail) had an almost complete agreement, with only one subject being classified differently (McNemar's test p = 1.000).

Eighty-one (77%) subjects failed the Optec 900 lantern while 24 (23%) subjects passed.

CIE 143:2001 classified 11 subjects as CIE 1, 16 in CIE 2 and 48 as CIE 3 (Table 3). The remaining 30 subjects were in the CIE 4 non-classifiable group, indicating that they have poor colour discrimination. The modified CAD CV severity classification found ten subjects to be CV 1 Normal, 11 CV 2 Safe, 41 CV 3 Poor and 43 who were CV 4 Severe.

According to HHR4 1-4 severity classification, there were ten CVN, 30 Mild CVD, 51 Moderate CVD and 14 Strong CVD. The weighted, HRR-w 1-4 differed by classifying ten as CVN, 35 as Mild, 50 as Moderate and ten as Strong CVD.

Table 3. Colour vision deficiency severity by CIE 143:2001, CAD CV, HRR and weighted HRR (HRR-w). In HRR-w, each plate receives a specific value, and the lowest score of either deutan or protan indicated the severity of the CVD.

	Colour vision classification			
Colour vision quality	CIE	CAD CV	HRR	HRR-w
I (Normal)	11	10	10	10
II (Mild)	16	11	30	35
III (Poor)	48	41	51	50
IV (Severe)	30	43	14	10
Total	105	105	105	105

The collapsed CAD CV 1-4 was found to correlate with CAD CV 1-5 indicating interchangeable use ($\tau = 0.97$, $p < 0.001$). Equivalently CIE 143:2001 correlated with CAD CV 1-4 ($\tau = 0.81$, $p < 0.001$), HRR ($\tau = 0.68$, $p < 0.001$) and HRR-w severity classification ($\tau = 0.70$, $p < 0.001$). The severity classifications were also compared after splitting the cohort according to CVD diagnosis.

A receiver operating characteristic (ROC) curve of Optec 900 (pass/fail) against CAD RG (continuous) gave an AUC of 0.96 (95 % CI 0.91–1.00); At 2.93 SN there was a specificity of 1.00 and sensitivity of 0.50, suggesting that all CVN subjects scored ≤2.93 SN. At 5.59 SN, all CVD subjects accurately failed (specificity = 0.93, sensitivity = 1.00) (Fig. 2). For an ROC of Optec 900 (pass/fail) against Ishihara's test (continuous number of incorrect), the AUC was 0.90 (95 % CI 0.83–0.99). Four incorrect responses on Ishihara's test gave a specificity of 1.00 and a sensitivity of 0.50. A ROC of FD15 (pass/fail) against CAD RG (continuous) gave an AUC of 0.98 (95 % CI 0.95–1.00) (Fig. 3). At 18.64 SN there was a specificity of 0.78 and a sensitivity of 1.00, which indicates that all CVD subjects scoring ≥18.64 SN will accurately fail FD15. The optimal Youden index for FD15 is at 14.40 SN (Specificity 0.94, Sensitivity 0.92).

Discussion

Comparing CIE 143:2001 to CAD colour vision

In this study, we found the strongest correlation ($\tau = 0.81$, $p < 0.001$) between CIE 143:2001 and CAD CV severity classification in the overall population, indicating that CAD CV may readily replace the CIE 143:2001. When comparing the subgroups of CVD, we found a strong correlation ($\tau = 0.77$, $p < 0.001$) for deutan, but a less strong correlation for protans ($\tau = 0.68$, $p < 0.001$). Use of the CAD CV severity classification instead of CIE 143:2001 will affect some subjects and their possibility to qualify for work. For the CVN, there was total agreement between the classifications. One mild deutan (CV 2; SN 2.90) was classified as CIE 1 since he passed Ishihara's test with two errors (Tables 3 & 4). However, the CAD CV Safe limit of 2.35 SN remains debatable. In CIE 1, HW-B is the lantern for a secondary test of CV. This threshold limit (2.35 SN) was set after a study of 226 deutan subjects by Barbur and Rodriguez-Carmona (2017) followed by a calculation of the correlation[2] between HW-A and HW-B; the so-called HW-A equivalence (Carter & Barbur 2015). The HW-B lantern is known to fail 10 % of CVN and all CVD, and a mild CVD is believed to be more likely to pass Ishihara's test than to pass HW-B. The confidence limits of the HW-A equivalence are not known, and we cannot rule out that this deutan subject is within the limits of safe job performance. Also, according to the ROC analysis of the Optec lantern test, a CAD score of 2.93 SN gave a specificity of 100%, indicating this as the limit were all CV Safe would pass.

Furthermore, this one mild deutan who was classified as CIE 1 would not be subject to a secondary test if the Ishihara's test were performed according to instructions before he was referred to us (Table 4). To add confidence in the CIE 1 classification, we suggest to implement HRR in addition

[2]HW-B (% error score) =27.503 + 0.649*HW-A (r2=0.72)

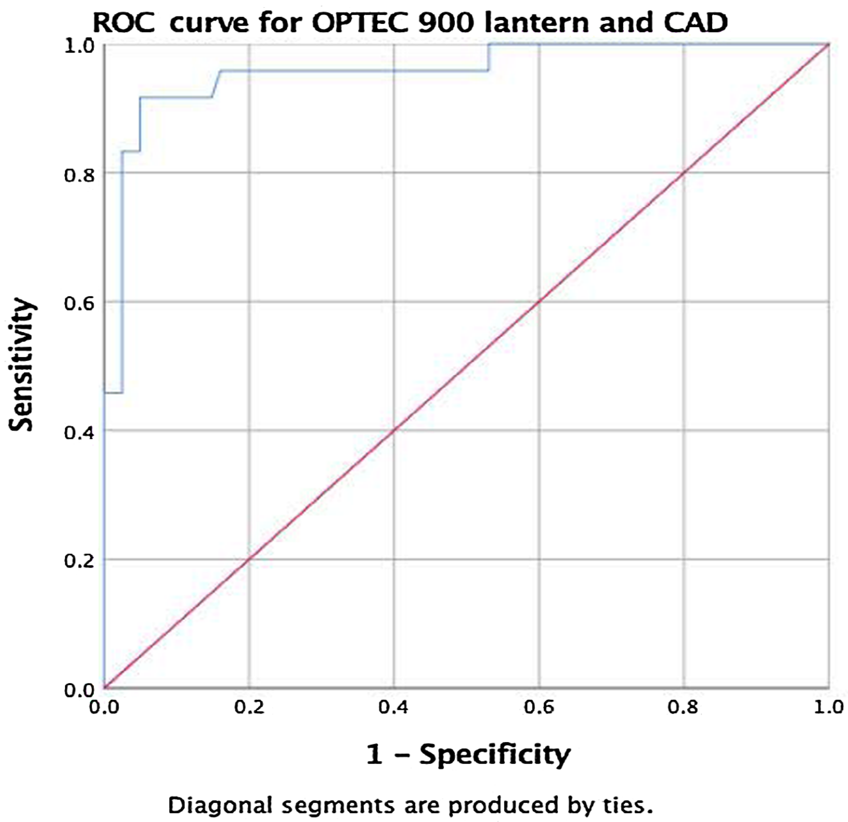

Fig. 2. Receiver operating characteristic (ROC) curve of a pass and fail of Optec 900 lantern with reference to CAD score. Twenty-four subjects passed the lantern test, and 81 failed. The area under the curve is 0.96 (95 % CI 0.91-1.00).

to Ishihara's test. It is unlikely that a CVD subject will pass both tests.

When discussing the CIE 2 classification, it is necessary to divide the subjects into deutan and protan CVD. Among 16 subjects classified as CIE 2, ten deutan subjects were classified as CV 2 safe and three CV 3 Poor by CAD. The three CV 3 Poor had CAD score of 4.11–5.00 but were classified as CIE 2 due to passing the OPTEC 900 lantern, the approved secondary test for CIE 2. This lantern is a Green A lantern and is built to recognize deutan and protan aviators with CVD within the range for safe flight (Holmes & Wright 1982). The CAD threshold limits, 6 SN for deutan and 12 SN for protan CVD, for safe flight was confirmed in a study on the correct identification of landing signal lights (Safety Regulation Group 2009). Again, the ROC for Optec gave the highest Youden's index (0.87) at 5.59 SN, which was consistent with the CAD deutan threshold for safe flight. The threshold limit of 12 SN for protans has not been validated for transport industries other than civilian aviation. Hence, this threshold limit cannot be applied for maritime lookout duties.

The classification of CV 3 (≤4.00 SN) is statistically equivalent to a pass of HW-A for all CVN, but this threshold does not match the specification of CIE 2 nor the Optec 900 lantern. Likewise, one of the protan CVD (7.03 SN) passed the OPTEC 900 lantern, which is consistent with the specification of the lantern and the threshold of 12 SN for protan CVD. The other protan in CIE 2, is classified as CV 4 Severe (13.12 SN), just above the expected threshold for passing the lantern test but just within the limit according to ROC of Optec (100% sensitivity at 13.26 SN). The CAA report questioned the 12 SN limit for protan CVD as seven of 27 with a threshold above 12 SN had correct identification of signal lights (Safety Regulation Group 2009).

The most discrepancy in categorization is for CIE 3 and CV 3 and 4 for the deutan subgroup. Thirty-five deutans and 14 protans classify as CIE 3. However, among deutans, 23 subjects classify as CV 3 Poor and 12 subjects as CV 4 Severe CVD. The main explanation for this is the cut-off level for the classifications. The qualification for CIE 3 is a pass of FD15 that is, with some limitations (Dain SA et al. 2019), consistent with the ability to recognize colour codes at a short distance. In this study, ROC analysis of FD15 found a sensitivity of 1.00 at 18.64 SN, which is consistent with the suggestion by Marechal & al. (2018) to classify CVD above 18.5 SN as dichromat.

In our study, the interval between CV 4 Severe (>12 SN) and 18.5 SN contain 13 deutan subjects (18%) who were classified as CIE 3 and one protan who was classified as CIE 2. Barbur and Rodrigues-Carmona (2017) estimate that 46% of deutan and 70% of protan CVD falls into category CV 4. Our data confirms that this is not a heterogenic group for colour vision performance. There are good reasons to diversify the CV 4 group into those who can recognize coloured surfaces and those who cannot. There was far less divergence between CIE 4 and CV 4 Severe. Only one deutan was classified as CIE 4 and CV 3 Poor. The Maritime and Coastguard Agency (MCA) has adopted CV 1 as the minimum standard for lookout duties and require a CAD test if Ishihara's test is failed. They do not include the CIE 2 and this will reduce the number of mild CVD applicants that may be considered fit for lookout duty in other nations. The MCA regulation is compliant with CIE 143:2001 for personnel without lookout duties, requiring an FD15 test.

Comparing CIE 143:2001 to HRR

When comparing CIE 143:2001 against HRR severity classification, we found a correlation of $\tau = 0.68$ ($p < 0.001$) for the combined CVDs. The deutans correlated $\tau = 0.59$ ($p < 0.001$), which was mainly because of divergence in the CIE 3 and CIE 4. In CIE 3, four subjects were classified as Mild CVD, 28 as Moderate CVD and three as Strong CVD. The HRR identified only nine of the protans identified by CAD, and this reduces the possibility to rely on the statistics for this subgroup.

The identification of protan is an essential part of CIE 143:2001, and this is challenging since the Medmont test is no longer available. To our knowledge, the HRR 4th edition is the only

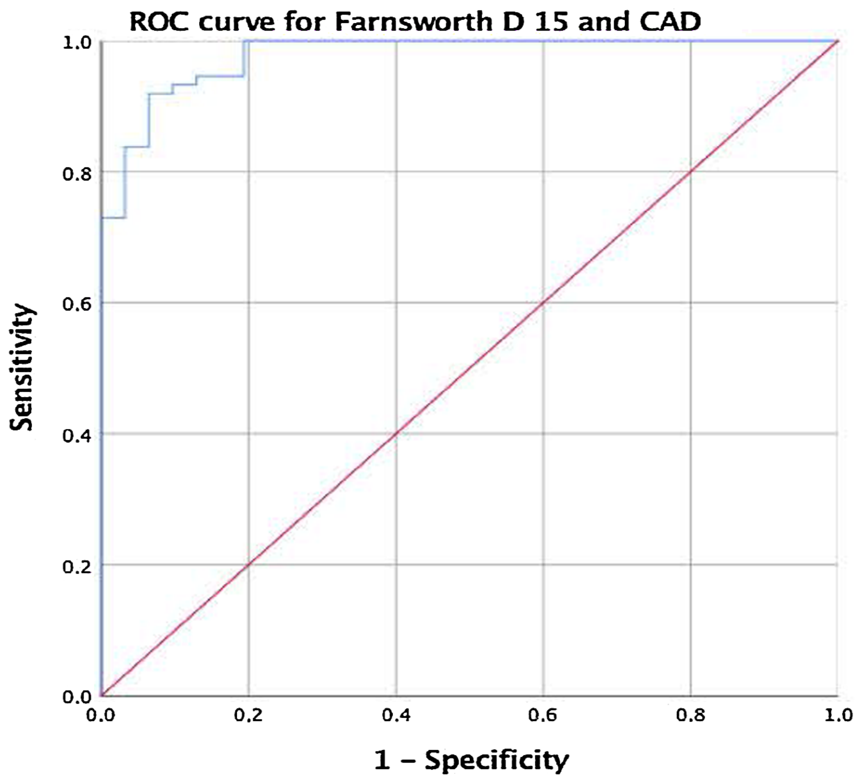

Fig. 3. Receiver operating characteristic (ROC) curve of a pass and fail of Farnsworth D15 with reference to CAD score. Seventy-four subjects passed the FD15 test, and 31 failed. The area under the curve is 0.98 (95 % CI 0.95-1.00).

Table 4. The severity classification, according to the CAD CV grading system and CIE 143:2001, stratified by CAD diagnosis. CAD CV classification is more conservative than CIE classification that uses Ishihara's test, Optec lantern and Farnsworth D15. The classification correlation for the deutan is τ=0.77 ($p < 0.001$), and for protan τ=0.68 ($p < 0.001$).

			CIE class				
CAD diagnose			CIE 1	CIE 2	CIE 3	CIE 4	Total
Normal		CV1 Normal	10				10
	Total		10				10
Deutan		CV2 Safe	1	10	0	0	11
		CV3 Poor	0	3	23	1	27
		CV4 Severe	0	0	12	21	33
	Total		1	13	35	22	71
Protan		CV3 Poor		2	12	0	14
		CV4 Severe		1	1	8	10
	Total			2	14	8	24
Total		CV1 Normal	10	0	0	0	10
		CV2 Safe	1	10	0	0	11
		CV3 Poor	0	4	36	1	41
		CV4 Severe	0	1	13	29	43
	Total		11	16	48	30	105

available PIP test that can identify the CVD subtypes and quantify the condition. However, we found HRR to misclassify 13 protans as deutan out of 24 protans and unable to classify two subjects. Bailey et al. (2004) make a point that the less saturated plates in HRR 4th edition are less able to differentiate protan and deutan. They suggest to assign a diagnostic value to each plate to increase the diagnostic accuracy. We did this modification to our data and were able to classify the two undiagnosed subjects. The result (HRR-w) was 11 deutan and 13 protan subjects out of the 24 protans diagnosed by CAD. The frequency of Mild CVD (15 individuals) was the same in HRR and HRR-w. As stated by Bailey & al. (2004), the greatest ambiguity in qualitative diagnosis is for the subjects with milder CVD. There was no significant difference in the severity of CVD for the deutan (11 subjects), and protan (four subjects) score for the subjects classified as Mild CVD in the HRR-w severity classification (Fig. 4). This finding implies that HRR is of little use for diagnosing the Mild CVD and that it has limited ability to substitute the Medmont test in the CIE 143:2001. However, in the HRR Moderate and Strong CVD, there were significant differences in diagnostic classification in HRR-w score, and this implies that HRR can be used to identify moderate and strong protans.

The idea behind the Medmont test was to be able to differentiate between protans and deutans, and a study by Metha and Vingrys (1992) indicated that the Medmont test is excellent for dichotomizing protans from deutans. The threshold for diagnostic use-values is <-2 for protans. In the CIE 143:2001 protocol, the cut-off for the Medmont test is <-1, which indicates a very conservative approach for approval of protans. The protocol states that it is necessary to identify protans when recognition of distant red lights is essential for safe operation. The lantern is considered a test of colour recognition and not a test for reduced sensitivity to red lights (CIE 2001).

Discussion of CIE 143:2001

The arguments referenced in CIE 143:2001 are all from before 2000 and do not reflect current knowledge of the genetics of CV nor new methods of CV testing. The quality of the results presented in several of the articles has severe limitations. In one central reference, Cole and Vingrys (1983) discuss the disadvantages for protans based on a population size of only five in each CV group -colour vision normal, deuteranope, protanope and protanomalous. In another study, Kinney & al. (1979), report the ability for CVD to judge light signals at sea. The study is indeed exciting but should be updated to reflect current diagnostic possibilities. An expert workshop in Kobe (2014) revised the CIE 143:2001

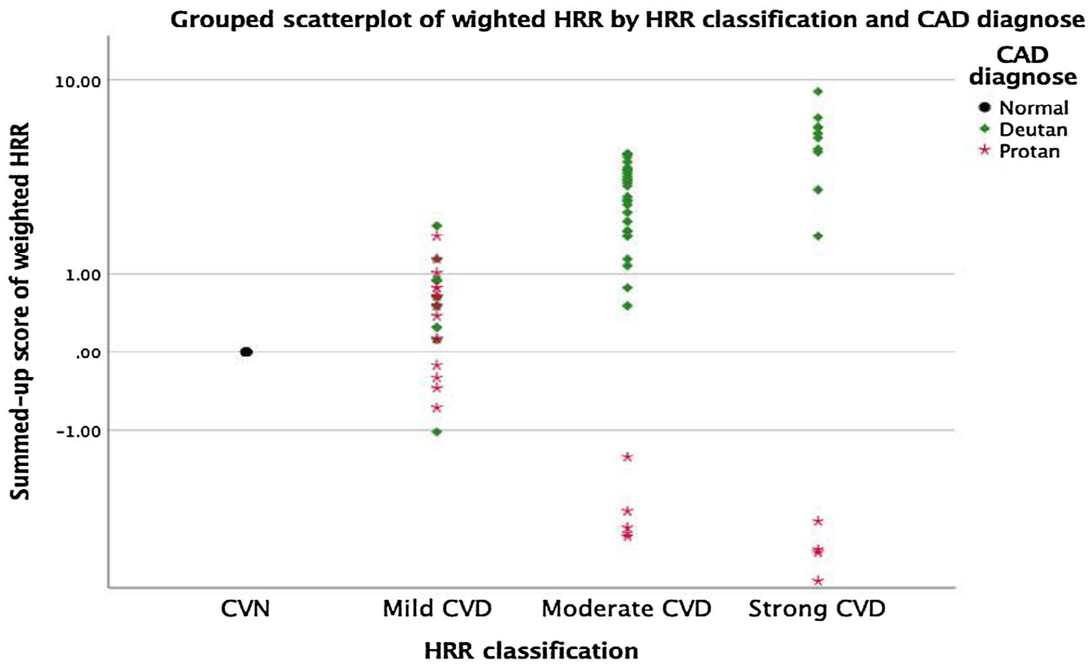

Fig. 4. Scatterplot of the summed-up score of weighted HRR (HRR-w), CV severity classification by HRR, and CV diagnosis by CAD. An HRR-w score of zero indicates a normal colour vision. A positive score indicates deutan CVD and a negative score indicates protan CVD. The protan CVD that scored positive were misclassified as deutan by HRR.

Table 5. New criteria for CIE colour vision classification using the adjusted tests for CIE 143:2001 or CAD test. The correlation for the two classifications is τ=0.90 ($p < 0.001$) for all diagnosis.

CAD diagnose			CIE class				
			CIE 1	CIE 2	CIE 3	CIE 4	Total
Normal		CV1 Normal	10				10
	Total		10				10
Deutan		CV2 Safe	1	10	0	0	11
		CV3 Poor	0	3	23	1	27
		CV4 Severe	0	0	12	21	33
	Total		1	13	35	22	71
Protan		CV3 Poor		2	12	0	14
		CV4 Severe		1	1	8	10
	Total			2	14	8	24
Total		CV1 Normal	10	0	0	0	10
		CV2 Safe	1	10	0	0	11
		CV3 Poor	0	4	36	1	41
		CV4 Severe	0	1	13	29	43
	Total		11	16	48	30	105

and stated in a report to CIE that none of the tests listed in CIE 143:2001 reliably quantifies the level of CVD. Also, there are no recent studies on visual requirements for safe maritime lookout duties or for other maritime work where colour discrimination is needed. We believe our study supports the critique of the tests listed in CIE 143:2001. The CIE 143:2001 CV severity classification is, to our opinion well justified, but the test methods and interpretations need to be updated.

Recommendations

We recommend the following tests for CIE 143:2001 severity classification (Table 5). Ishihara's test can still classify CIE 1 as the initial test. Ishihara's test has high positive and negative predictive value if performed correctly. An error-score ≤2 errors on plate 2–16 is the preferred choice with a test accuracy of 0.99 based on the results published by Barbur and Rodrigues-Carmona (2017). The accuracy can be improved with the addition of HRR in the classification of CIE 1.

The alternative test will be CAD. CIE 1 and 2 qualify for duty as a lookout as do CV 1 Normal. CV 2 Mild is reasonably consistent with CIE 2 in our study, but do not qualify for navigational lookout duties in the United Kingdom (MCA 2014). UK has adopted the limit of 2.35 SN based on a theoretical approach. Although our data indicate 2.93 SN as the limit, but we support the 2.35 SN threshold limit until adequate visual task analysis has been performed to justify other CV requirements. We also support keeping CIE 2, as the personnel classified in this group do not exhibit normal CV, and must be expected to have reduced capacity to perform colour critical tasks. The Medmont test is no longer available, and there is no other functional substitute to diagnose CVD in this group of mild to moderate CVD except for CAD. Job performance analyses of aviators (Safety Regulation Group 2009) find safe limits of 6 SN for deutans and 12 SN for protans. These limits are consistent with our

Table 6. The severity classification, according to the new CAD threshold grading system and CIE 143:2001, stratified by CAD diagnosis. New CIE 1 ≤ 2.35 SN, CIE 2 2.36-6.0 SN, CIE 3 6.01-14.40 SN and CIE 4 > 14.40 SN. The classification using CAD is more conservative than by CIE classification using Ishihara's test, Optec lantern and Farnsworth D15. The classification correlation for the deutan is τ=0.87 (*p* < 0.001), and for protan τ=0.84 (*p* < 0.001).

CAD diagnose			CIE class				
			CIE 1	CIE 2	CIE 3	CIE 4	Total
Normal		CAD CIE 1	10				10
	Total		10				10
Deutan		CAD CIE 2	1	13	2		16
		CAD CIE 3			28	1	29
		CAD CIE 4			5	21	26
	Total		1	13	35	22	71
Protan		CAD CIE 3		2	13	0	15
		CAD CIE 4			1	8	9
	Total			2	14	8	24
Total		CAD CIE 1	10				10
		CAD CIE 2	1	13	2		16
		CAD CIE 3		2	41	1	44
		CAD CIE 4			6	29	35
	Total		11	15	49	30	105

results on the Optec 900 lantern, but cannot readily be transformed into use in maritime fitness standards.

Furthermore, CIE 3 classifies everyone that is considered safe on identifying colour surfaces. On Farnsworth D15, ≤2 diametrical crossing will best correspond to 14.40 SN. Individuals failing FD15 or who score above 14.40 must be considered to have severe CVD. The correlation for the two CIE 143:2001 severity classifications is τ = 0.90 (p < 0.001) for all diagnosis (Table 6).

Our suggestion for CIE 143:2001 and the implementation of CAD is an intermediate solution to facilitate continued use of the standard, until scientifically valid and up to date studies of visual requirements for safe work-performance are performed.

Conclusion

The colour vision severity classification standard 'CIE 143:2001 International recommendations for colour vision requirements in transport', has become out of date due to lack of commercial availability of required colour vision tests. A direct link between the methods for evaluating colour vision and the colour visual performance in transport is still poorly funded and will require task performance studies.

New diagnostic tools with better accuracy have not been implemented in the standard. Meanwhile, we propose three possible revisions to the CIE 143:2001 algorithm, based on the availability of CAD:

1 Replacing the current CIE 143:2001 algorithm using new CAD threshold limits
2 Use of CAD as a secondary test to Ishihara's test and HRR or
3 Revising the current CIE 143:2001 algorithm using Ishihara's test, HRR, Optec 900 and FD15

When using 'Colour Assessment and Diagnosis' (CAD) as a replacement for the current CIE 143:2001 algorithm, we suggest new threshold limits. The recommended limits are CIE 1 (normal) ≤2.35 SN, CIE 2 (mild CVD) (2.36–6.0 SN, CIE 3 (severe CVD) 6.01–14.40 SN and CIE 4 (non-classifiable CVD) > 14.40 SN.

Because a CVD is unlikely to pass both Ishihara's test and HRR, these two tests will accurately differentiate CVD and CVN (CIE 1 normal) subjects. Those who fail these tests can be further examined and classified using CAD as a secondary test.

In the absence of CAD, the severity classification can be done using a pass on Ishihara's test and Hardy Rand Rittler (HRR) 4th edition for CIE 1, Optec 900 lantern for CIE 2 and Farnsworth D 15 for CIE 3 colour vision standard. Those who fail on all these tests would be considered as non-classifiable CVD (CIE 4).

References

Maritime and Coastguard Agency (MCA). (2014): Guidance: The approved doctor's manual. https://www.gov.uk/government/publications/the-approved-doctors-manual.

Bailey JE, Neitz M, Tait DM & Neitz J (2004): Evaluation of an updated HRR color vision test. Vis Neurosci **21**: 431–6.

Bailey KGH & Carter T (2016): Consistency of secondary colour vision tests in transport industries. Occup Med (Lond) **66**: 268–75.

Baratloo A, Hosseini M, Negida A & El Ashal G (2015): Part 1: simple definition and calculation of accuracy, sensitivity and specificity. Emergency (Tehran, Iran) **3**: 48–9.

Barbur JL & Rodriguez-Carmona M (2017): Colour vision requirements in visually demanding occupations. Br Med Bull **122**: 51–77.

Carter T & Barbur J (2015): Colour vision assessment for maritime navigation lookout. Review for UK Maritime and Coastguard Agency (MCA). Maritime and Coastguard Agency.

CIE. (2001): International recommendations for colour vision requirements in transport. Technical report 143:2001. Commision Internationale de L'Eclairage.

CIE. (2019): Calometry - Part 1: CIE Standard Calometric Observers. Commision Internationale de L'Eclairage.

Cole BL, Vingrys AJ(1983): Do protanomals have difficulty seeing red lights. Proc CIE 20th Session, Amsterdam, 1983. Guideline 56, E04, CIE Paris, 1, 3

Dain SJ, Atchison DA & Hovis JK (2019): Limitations and precautions in the use of the farnsworth-munsell dichotomous D-15 test. Optomet Vis Sci **96**: 695–705.

Expert workshop, University of Kobe. (2014): Test Methods for Color Vision in Seafarers with Navigational look-out duties. Edited by International Maritime Health Association.

French A, Rose K, Cornell E & Thompson K (2008): 'The evolution of colour vision testing. Aust Orthoptic J **40**: 7.

Holmes JG & Wright WD (1982): A new color-perception lantern. Color Res Application **7**: 82–88.

Holmgren F (1878): Colour-blindness in its relation to accidents by rail and sea.

IALA. (2017): R0201 (E-200-1) marine signal lights – colours. International Association of Marine Aids to Navigation and Lighthouse Authorities.

Kinney JAS, Paulson HM & Beare AN (1979): The ability of color defectives to judge signal lights at sea. J Opt Soc Am **69**: 106–13.

Marechal M, Delbarre M, Tesson J, Lacambre C, Lefebvre H & Froussart-Maille F (2018): Color vision tests in pilots' medical assessments. Aerospace Med Human Perform **89**: 737–43.

Metha AB & Vingrys AJ (1992): The C-100: A new dichotomiser of colour vision defectives. Clin Exp Optometry **75**: 114–23.

Mollon JD & Cavonius LR (2012): The Lagerlunda collision and the introduction of color vision testing. Surv Ophthalmol **57**: 178–94.

Safety Regulation Group (2009): Minimum colour vision requirements for professional flight crew. Recommendations for new colour vision standards. CAA paper. Federal Aviation Administration.

Vingrys AJ & Cole BL (1986): Origins of color-vision standards within the transport industry. Ophthal Physiol Opt **6**: 369–75.

Wilson G (1855): Researches in colour-blindness: With a supplement on the danger attending the present system of railway and marine coloured signals 1818–1859. Edinburgh: Sutherland & Know.

Received on January 18th, 2020.
Accepted on March 30th, 2020.

Correspondence:
Vilhelm F. Koefoed, MD, PhD
Department of Global Public Health and Primary Care
University of Bergen
5020 Bergen
Norway
Tel: +4790727527
Fax: +4755586130
Email: v@koefoed.no