

THE UPTAKE AND PERFORMANCE OF WHO 5TH EDITION MANUAL METHODS TO ASSESS SPERM MORPHOLOGY SINCE 2010: EXPERIENCE FROM AN AUSTRALIAN EXTERNAL QUALITY ASSURANCE (EQA) PROGRAMME.

Michelle Kitson¹, Phillip Matson², and Emily Zuvela²

¹School of Medical and Health Sciences, Edith Cowan University, Joondalup, and ²External Quality Assurance Schemes for Reproductive Medicine, PO Box 162, Northlands, Australia.

INTRODUCTION AND AIMS

The assessment of sperm morphology is subjective and different classification systems have been used around the world over time. In an attempt to standardize all methods used in the analysis of semen, the World Health Organization (WHO) produced its first manual in 1980, revising it in 1987, 1992 and 1999, culminating in the current version 5th Edition (WHO5) produced in 2010. Interestingly, a draft of the 6th Edition has just been sent out for public review.

The classification of normal sperm has varied from the early liberal methods whereby MacLeod and Gold (1951) were "not prepared at this time to classify any but the most distorted forms as truly "abnormal" cells in the sense that we deny the ability of any particular cell to fertilize the ovum." This contrasts with the WHO5 that states that we should "limit what is identified as normal to the potentially fertilizing subpopulation of spermatozoa prevalent in endocervical mucus". The method recommended in WHO5 is a simple normal/abnormal classification, and all borderline forms are to be considered abnormal

The aims of the present study were to (i) identify changes in morphological criteria used by clinical laboratories since before the release of WHO5, (ii) determine the consistency with which laboratories identify sperm with normal morphology since the release of WHO5, and (iii) assess the between-laboratory precision over time.

MATERIALS AND METHODS

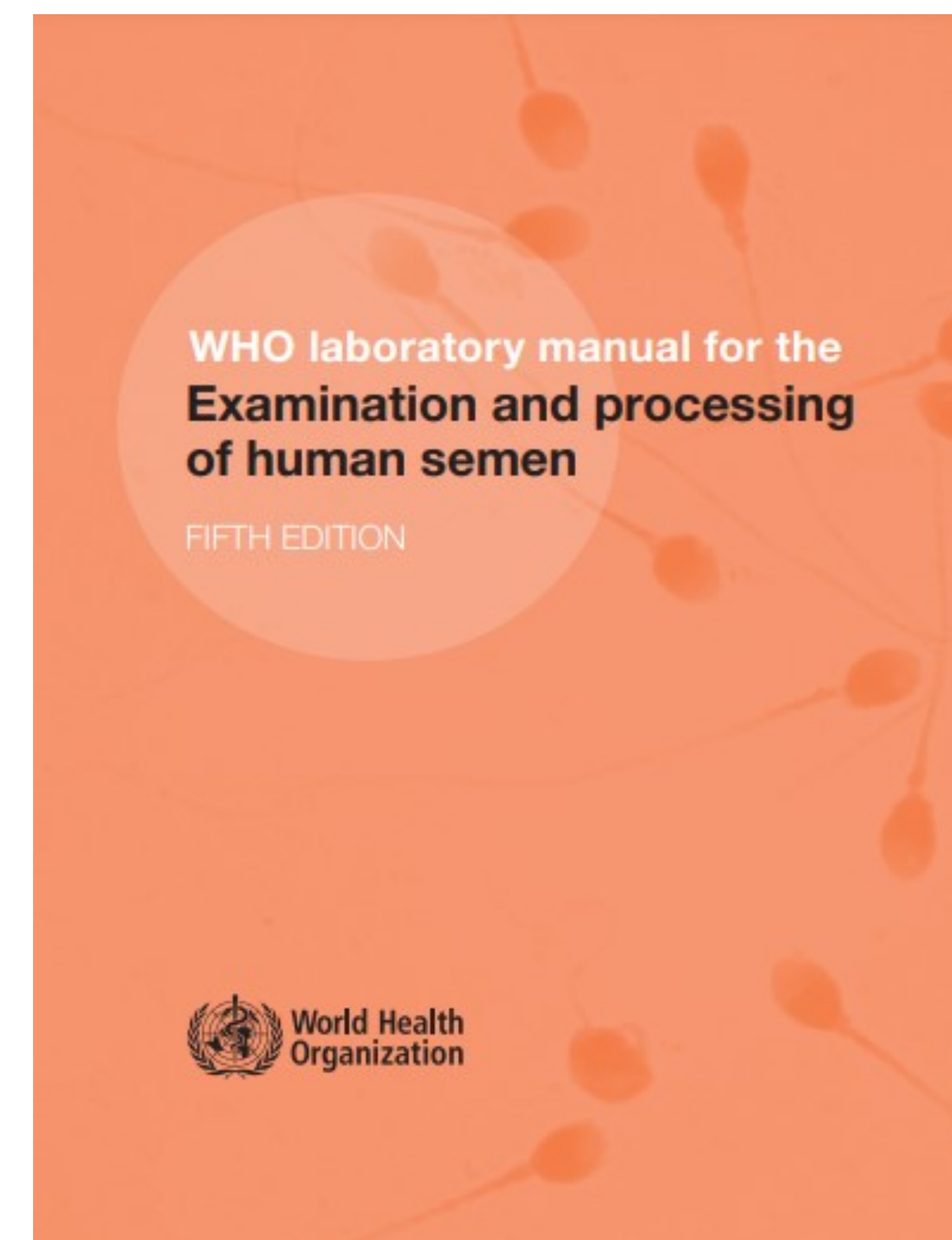
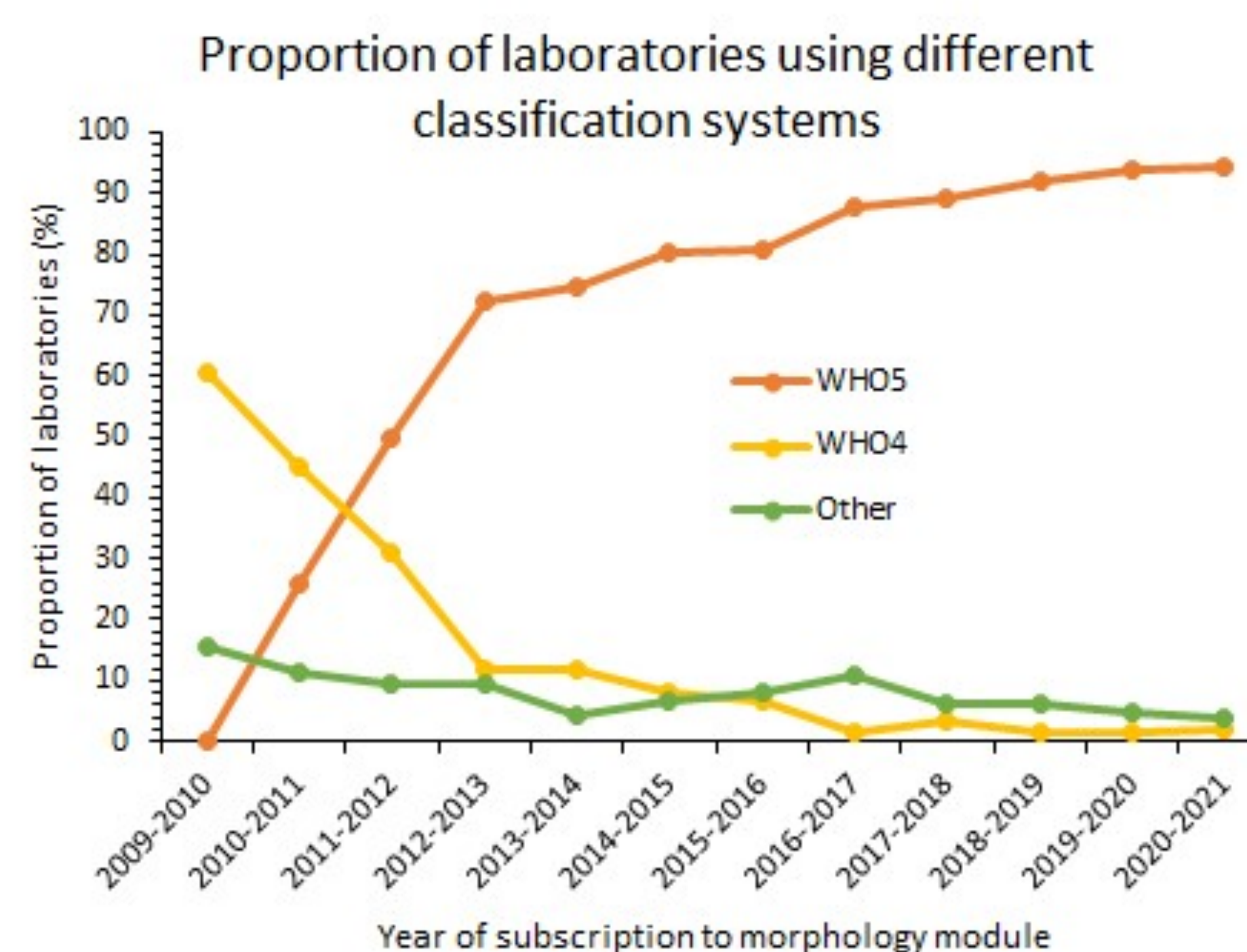
Samples were posted to laboratories enrolled in the sperm morphology module. Distributions were made every three months according to a pre-determined schedule, and each distribution included 3 'wet' semen samples and a slide with a semen smear stained with Diff Quik. Results were submitted online by laboratories via the EQASRM website within five weeks of sample dispatch, and statistical summaries of the results were then available for download after the closing date. In addition to the proportion of sperm with normal morphology, laboratories were asked to state the lower reference limit used when reporting clinical samples, and these then linked to the recognised morphology classification system used. Summaries of the results for the pre-stained slides were excluded from this study to ensure the laboratories were analysing samples stained under their routine protocols.

Laboratories included in the study used a lower reference limit of 4% normal forms, and this is described in the WHO 5th Edition manual. The classification of normal sperm morphology is said to align closely with the Strict criteria.

Laboratories were asked to report for each sample the proportion of sperm with normal morphology. The sample summaries supplied back to the laboratories contained the following information: (i) the all laboratory-trimmed mean (ALTM) and range for all laboratories, (ii) the overall mean and range for accredited ART laboratories, (iii) the method-related mean and standard deviation (SD), and (iv) the bias of the laboratory's own results from the method-related mean.

UPTAKE OF WHO5 METHODOLOGY

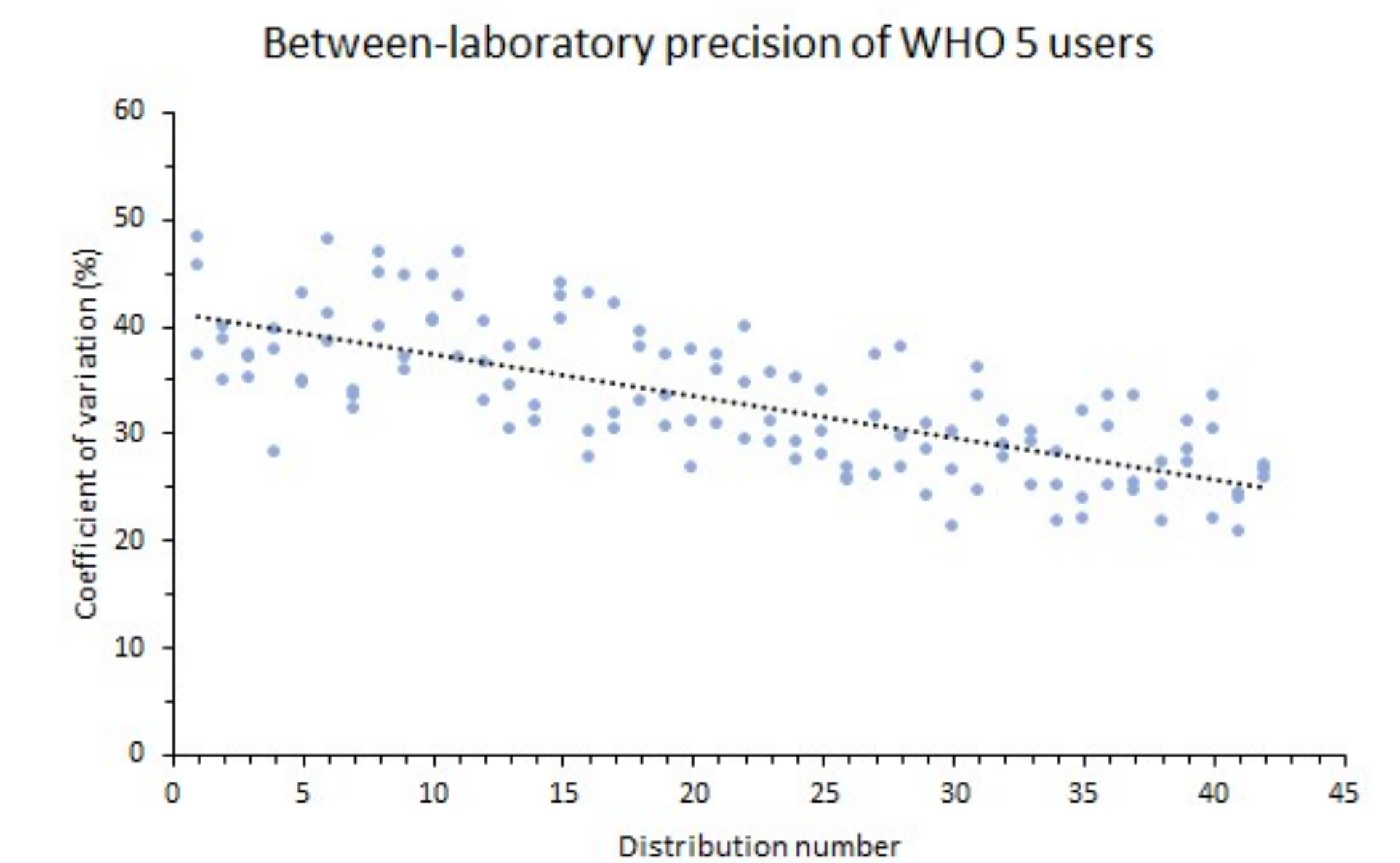
In 2009-2010, prior to the release of WHO5, there was a wide range of classification criteria in use including Strict criteria and those described in WHO manuals going back to 1987. However, the criteria defined in the WHO5 and using a lower reference limit of 4% normal forms were adopted by 50% of laboratories in the first two years following its initial publication and went up to 80% in five years, confirming a major effort by laboratories to update their protocols to use the most recent WHO manual. This was mainly at the expense of the use of WHO4 dropping from 61% to 8% in the five years since the introduction of WHO5. By 2018-2019, more than 90% of laboratories were using the one classification system of the WHO5.



BETWEEN-LABORATORY VARIATION FOR EACH SAMPLE

Between-laboratory variation was assessed for each sample by calculating the coefficient of variation as (standard deviation*100)/mean.

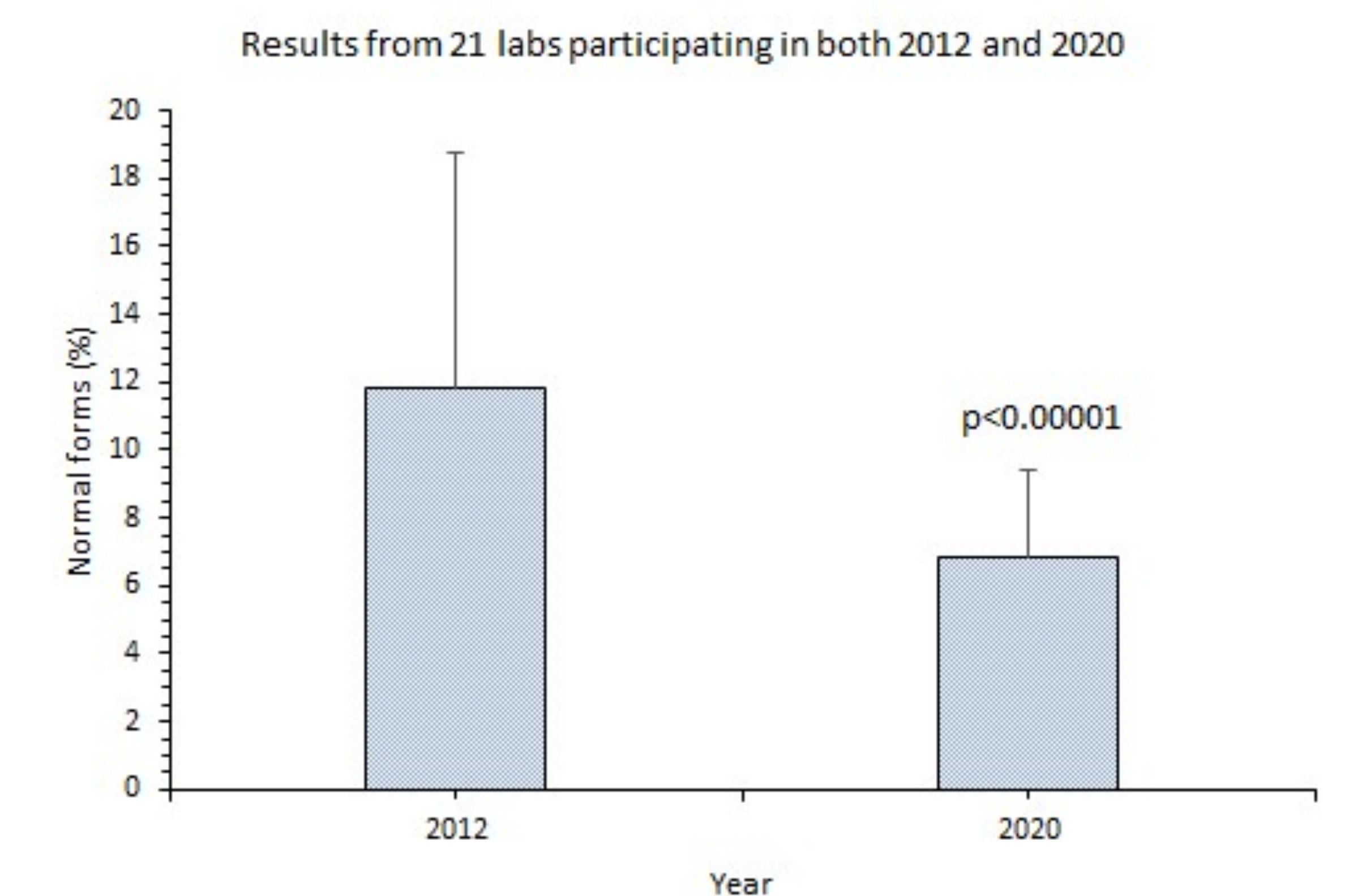
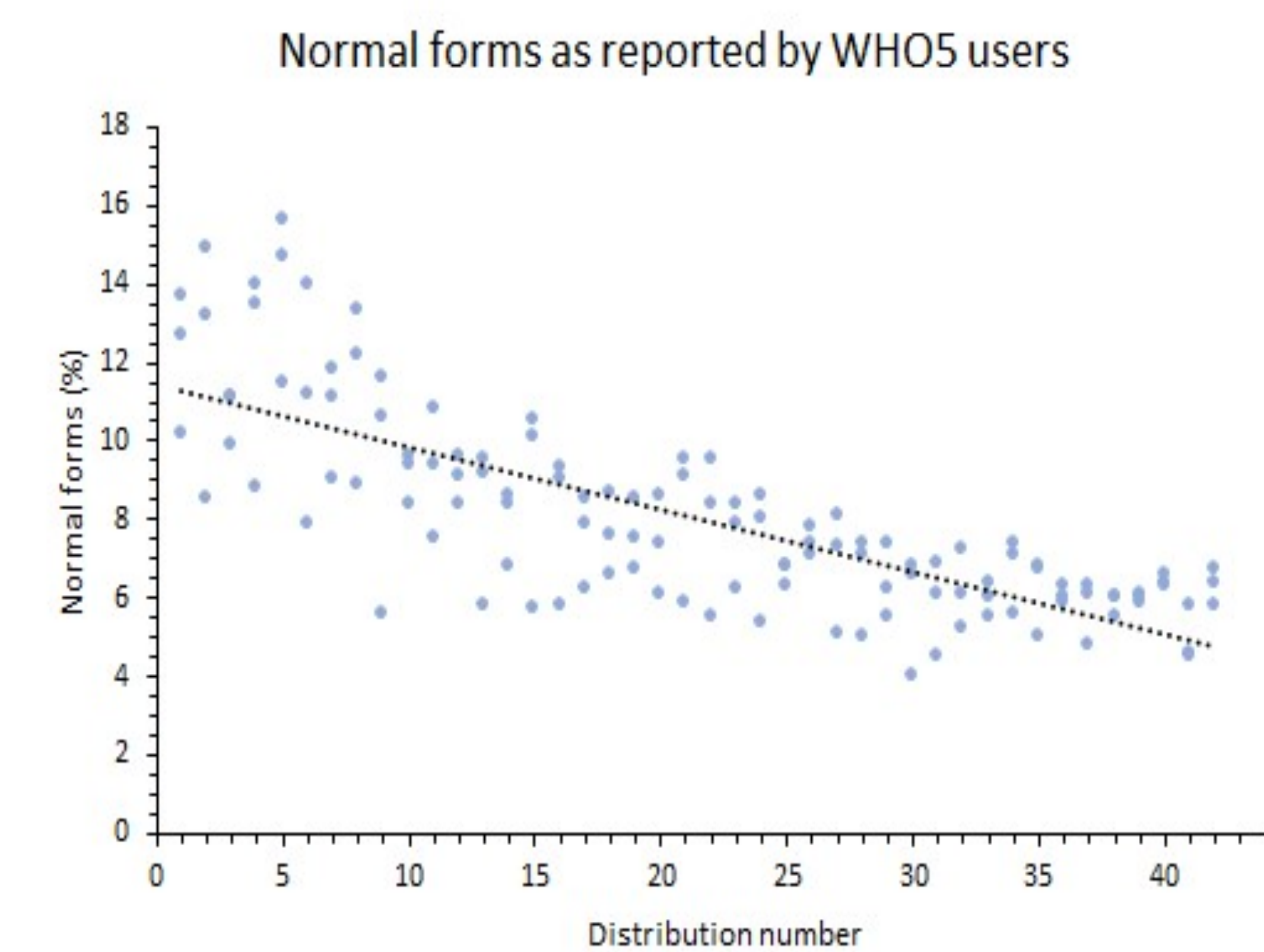
At the introduction of WHO5 in 2010-11 the between-laboratory variation was high with a coefficient of variation as high as 48%. As WHO5 continued to be used, there was a reduction in between laboratory with coefficients of variation reducing to as low as 27% in the 2020 – 2021 reporting periods. This improvement in between-laboratory reproducibility is reflected in a reduction at each distribution in coefficient of variation (%) of mean normal reported forms when plotted against distribution number ($r=-0.7279$, $p < 0.00001$).



INSTABILITY OVER TIME IN REPORTING PROPORTION OF NORMAL FORMS

In 2010-2011 when WHO5 was first introduced, the proportion of normal forms reported was well over 10%, (below, left). As more laboratories adopted the WHO5 classification methodology and became more experienced, the percentage of normals forms reported reduced over time ($r=0.7806$, $p<0.00001$). It appears that laboratories became more 'strict' in their identification of normal forms as time progressed, with the numbers of normal forms reported more than halved over the last ten years. This downward trend in normal forms reported, as well as the tighter clustering of values is clearly visible.

The reporting of normal forms (below, left) shows the average result per sample, this does not take into consideration the population of laboratories participating in each distribution. Therefore, a different approach was taken identifying 21 laboratories who were enrolled and submitted results in both 2012 and 2020. The results reported by these 21 labs in the two different years is shown below right. It shows that the results in 2012 were significantly higher than those reported in 2020, and this is consistent with the distribution summaries mentioned above.



SUMMARY

In summary, (i) the introduction of WHO5 resulted in the effective adoption of the its morphology classification system, (ii) laboratories showed improved between-laboratory variation over time, (iii) the identification of normal forms over time was inconsistent as laboratories became more strict, and (iv) given the reduction in reported normal forms over time when working from the same WHO5 manual, it appears that either increased training of laboratory personnel or the consideration of validated objective automated analysers would seem warranted.