Georg Hoffmann, Nina Allmeier, Modupe Kuti, Stefan Holdenrieder and Inga Trulson*

How Gaussian mixture modelling can help to verify reference intervals from laboratory data with a high proportion of pathological values

https://doi.org/10.1515/labmed-2024-0118
Received July 10, 2024; accepted August 1, 2024;
published online August 19, 2024

Abstract

Objectives: Although there are several indirect methods that can be used to verify reference limits, they have a common weakness in that they assume a low proportion of pathological values. This paper investigates whether a Gaussian decomposition algorithm can identify the non-pathological fraction even if it is not the main subset of mixed data.

Methods: All investigations are carried out in the R programming environment. The mclust package is used for Gaussian mixture modelling via the expectation maximization (EM) algorithm. For right-skewed distributions, logarithms of the original values are taken to approximate the Gaussian model. We use the Bayesian information criterion (BIC) for evaluation of the results. The reflimR and refineR packages serve as comparison procedures.

Results: We generate synthetic data mixtures with known normal distributions to demonstrate the feasibility and reliability of our approach. Application of the algorithm to real data from a Nigerian and a German population produces results, which help to interpret reference intervals of reflimR and refineR that are obviously too wide. In the first example, the mclust analysis of hemoglobin in Nigerian women supports the medical hypothesis that an anemia rate of more than 50 % leads to falsely low reference limits. Our algorithm proposes various scenarios based on the BIC values, one of which suggests reference limits that are close to published data for Nigeria but significantly lower than those established for the Caucasian population. In the second example, the standard statistical analysis of creatine kinase in German patients with predominantly cardiac diseases yields a reference interval that is clearly too wide. With mclust we identify overlapping fractions that explain this false result.

Conclusions: Gaussian mixture modelling does not replace standard methods for reference interval estimation but is a valuable adjunct when these methods produce discrepant or implausible results.

Keywords: machine learning; reference interval; Gaussian mixture modelling; reflimR; refineR

Introduction

According to current guidelines, it is mandatory that laboratories verify the reference limits obtained from external sources before using them for routine clinical care [1, 2]. This is especially true, when a laboratory introduces a new or changes an existent analytical method. For this purpose, the application of indirect methods based on retrospective data from the laboratory information system has been recommended [2, 3]. Although these methods are much more practicable than the direct procedures, which require carefully selected reference individuals [3, 4], they have a common weakness in that they assume a low proportion of pathological values.

In general, this requirement may be met by laboratory values from routine clinical tests, but there are certain situations where a high proportion of pathological values is unavoidable. These values may distort standard reference interval (RI) verification methods, especially if there is a large overlap between pathological and non-pathological results.

To alleviate this problem, we hypothesized that statistical learning methods for distribution mixture modelling should provide indications of the existence of strongly overlapping clusters and thus enable better estimation of reference intervals. Among the various solutions available in the R environment [5], we chose the mclust package [6] because it is able to automatically predict the best Gaussian mixture model for laboratory data. The key question is whether mathematically motivated fitting would lead to medically plausible models.

***Corresponding author: Inga Trulson**, German Heart Center Munich, Institute of Laboratory Medicine, Munich, Germany,
E-mail: trulson@dhm.mhn.de
Georg Hoffmann, Nina Allmeier and Stefan Holdenrieder, German Heart Center Munich, Institute of Laboratory Medicine, Munich, Germany
Modupe Kuti, College of Medicine, University of Ibadan, Ibadan, Nigeria; and Synlab Nigeria Laboratories, Lagos, Nigeria

Materials and methods

Data

In this study, we use both simulated and real-world laboratory data for establishing and validating our methodological approach. We create synthetic mixtures of normally distributed random samples with known reference limits by calculating the means and the standard deviations of the respective distributions from their 0.025 and 0.975 quantiles:

$$\text{mean} = (q_{0.025} + q_{0.975})/2$$

$$\text{sd} = (q_{0.975} - q_{0.025})/3.92$$

For lognormal distributions the same equations are applied to the logarithms of the quantiles.

Retrospective real-world data are retrieved from the laboratory information systems of Synlab Nigeria and the German Heart Center Munich. Ethical approval for the use of routine laboratory data has been approved by the National Health Research Ethics Committee (NHREC) in Nigeria (19/12/2008a; Health Research Committee assigned number: ADM/DSCST/HREC/APP/6420) and an ethics vote in respect of the use of routine data for quality control is given by the Technical University Munich (10/21 S-KH).

Data preparation

The data sets are filtered so that only the first (i.e. the oldest) value per patient is included in the analysis [7]. Apart from value, age and sex, the evaluation records do not contain any personal information about the patients. The Nigerian study encompasses 26,779 adult female outpatients (18–75 years of age) with no preselection by diagnosis, whereas the German study is conducted with 8,799 hospitalized men aged 55–80 years with a history of cardiac diseases such as heart failure or myocardial infarction. Patients with highly pathological values (hemoglobin <5 and >20 g/dL, CK>500 U/L) are excluded from this study to make the graphics easier to read. Values far outside the reference interval under consideration are removed by the standard procedures anyway and therefore do not influence the results.

Methods

To estimate reference limits with standard statistical methods, we apply the R packages reflimR [8] and refineR [9], and for the mixture deconvolution approach described here, we use the R package mclust [6]. In their simplest version, all three methods can be called with a single line of code, namely reflim(x), findRI(x) or Mclust(x) where x represents quantitative laboratory results to be analyzed for RI verification.

Theoretical part

The application of the new method is based on the assumption that distributions of laboratory values measured in healthy reference individuals are either symmetric or right-skewed and can be modeled either by a normal or a lognormal distribution [1, 10]. Left-skewed distributions are assumed to represent a mixture of normal distributions with a notable proportion of slightly decreased values.

The first group of analytes mainly includes blood constituents such as hemoglobin, sodium, albumin or glucose that fulfill a physiological function in blood and are therefore strictly controlled by the organism. The second group of analytes is much larger and comprises substances such as metabolites (e.g. bilirubin, creatinine) or compounds from damaged cells (e.g. ALT, CK), which are released into the blood stream without having a specific function there.

The distribution of values measured in the context of patient care is complex, due to the overlap of physiological and pathological processes. This forms the basis of our hypothesis that such distributions can be modeled by several overlapping normal (Gaussian) or lognormal distributions. If this assumption is correct, one of the distributions identified by the algorithm should reflect the healthy reference cohort, while the others represent patient cohorts, which have no significance for the verification of reference intervals.

The challenge of the new approach is visualized in Figure 1. It shows a mixture of simulated hemoglobin values resembling those measured in a population with a high prevalence of anemia. The dataset contains 500 normally distributed values with a target reference interval of 12–16 g/dL for women [11]. Another 500 values are composed of various normally distributed fractions with lower target values resulting in a seemingly homogenous left-skew of the total density curve. Compared to the expected limits, the reference interval estimated with indirect methods is too wide and shifted to the left (reflimR*:* 9.7–16.3 g/, refineR*:* 10.9–15.9 g/dL).

The crucial question is whether the machine learning algorithm can define a mixture of Gaussian distributions, which explains the left-skewed density curve mathematically and at the same time gives a medical explanation, why the standard methods yield incorrect reference intervals.

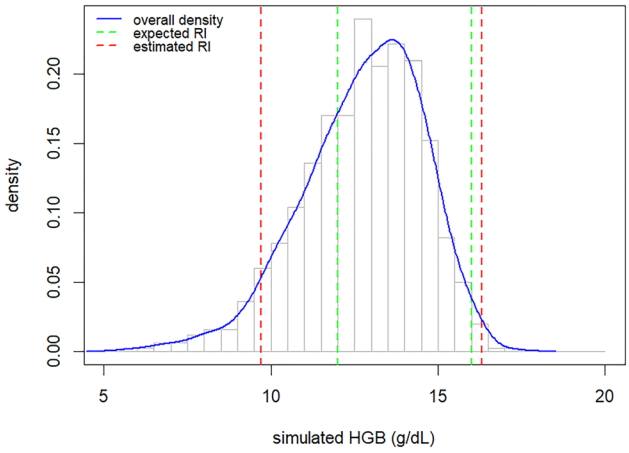

Figure 1: Distribution of hemoglobin (HGB) with expected and estimated reference limits with reflimR.

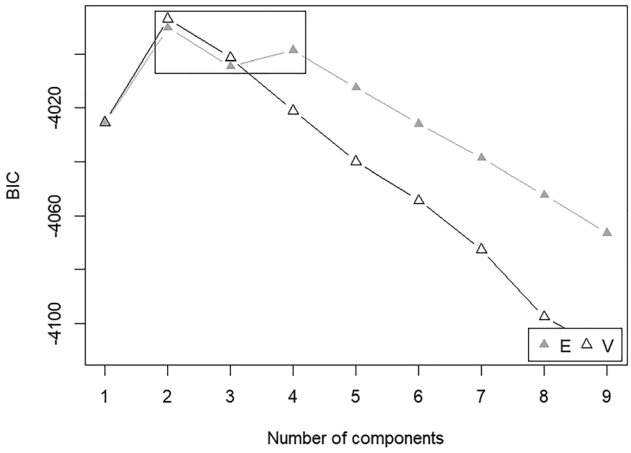

Figure 2: BIC values for a total of 18 models, nine with equal (E) and nine with variable (V) variance. The black rectangle outlines models with the highest Bayesian Information Criterion (BIC) values.

The mclust package used in this study is a popular and very powerful tool for model-based clustering, classification, and density estimation [6]. It generates Gaussian finite mixture models by fitting a range of distributions to the underlying data. The likelihood of an initial fit based on hierarchical clustering of the data is iteratively increased using the expectation maximization (EM) algorithm. Its goal is to find the parameter estimates (mean values and standard deviations) that maximize the log-likelihood of the observed data under the Gaussian mixture model. To this end, the two steps (expectation and maximization) are repeated until convergence, i.e. until the parameter estimates stabilize and the change in the log-likelihood becomes negligible [6].

The optimal number of clusters with either equal or variable variance is identified by the Bayesian Information Criterion (BIC)

$$\mathrm{BIC} = \log(L) - 0.5 \cdot p \cdot \log(n)$$

where log(L) is the log-likelihood of the model given the data x [6], p is the number of parameters in the model and n is the number of data points (measurements). A higher log-likelihood indicates better fit, whereas a higher number of parameters and/or data points reduces the BIC in the sense of a penalty term to avoid overfitting.

The model with the highest BIC value is selected automatically because it usually indicates the best trade-off between model fit and complexity. However, the arguments G and modelNames of the Mclust function can be changed manually if the investigator feels that a model with lower BIC fulfills medical assumptions better. For example, the function call Mclust (data=x, G=3, modelNames="V") will create a model of x with three clusters of variable variances (model V3).

Figure 2 illustrates the BIC analysis for the above hemoglobin data. Two clusters with either variable (V) or equal (E) variance, have nearly identical BIC values of –3,987 and –3,990, respectively. However, models with three and four clusters perform only slightly worse, with BIC values ranging from –3,999 to –4,005. This indicates that, from a statistical standpoint, models with three and four clusters fit almost as well (V3, E3, E4).

Figure 3 compares four top models with two to four clusters in terms of their shares in percent and reference limits in g/dL. The algorithm automatically selects model V2 (two clusters with variable standard deviation), reflecting about equally large pathological and non-pathological fractions. The reference interval of 11.6–16.0 g/dL comes much closer to the target values of 12–16 g/dL than those obtained with reflimR and refineR. Among the other models, V3 and E4 each contain a fraction with similarly fitting reference intervals, but E4 splits the non-pathological fraction into two almost identical clusters. This can be interpreted as a sign of overfitting.

Results

In the first part of this study, we investigate the question to what extent the statistical learning method described here is suitable for explaining unplausible results of the standard methods when the underlying sample contains a high percentage of pathological values. The statistical background of this choice is that hemoglobin values of healthy individuals can be modelled Gaussian, and the medical background of the experiment is the known prevalence of low hemoglobin results in the sub-Saharan African population [12].

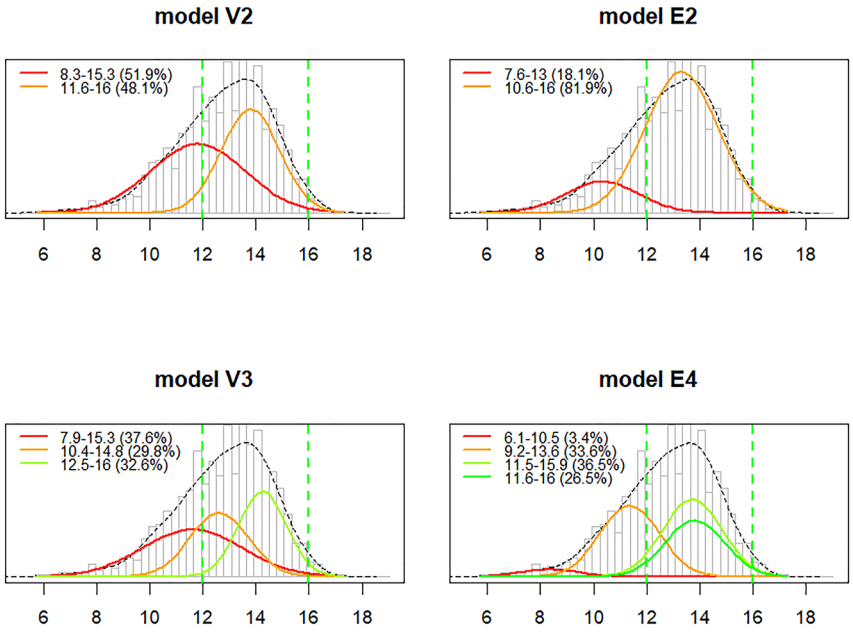

Figure 3: Illustration of four models with high BIC values around −4,000 (see box in Figure 2).

Hemoglobin

Figure 4 demonstrates that the hemoglobin concentrations measured in Nigerian women do not change very much throughout life so that there is no need for partitioning the data by age. Compared to the reference limits commonly used for the Caucasian population (about 12–16 g/dL) [11], the maximal density of the African measurements (dark region in Figure 4) is around the lower limit rather than in the middle of the reference interval for women.

The results of the analysis with reflimR and mclust are summarized in Figure 5. The reflimR algorithm yields a reference interval as wide as 8.1–14.5 g/dL. The red bars in the upper left graphic indicate that the divergence from the Caucasian target values significantly exceeds the medical tolerance range [8]. With 8.9–14.2 g/dL, the respective refineR estimate was comparable albeit a bit narrower (graphic not shown). Application of the mclust algorithm resolves the overall density into three largely overlapping clusters with variable standard deviation (model V3) that support the medical hypothesis of a high proportion of anemic women in Nigeria. One of these clusters with a share of about 30 % and a reference interval of 10.5–14.1 g/dl could represent the healthy female population in Nigeria, while the other two clusters with proportions of about 20 and 50 % would correspond to populations with mild and severe anemia, respectively (see discussion).

Although this suggestion of three clusters sounds plausible from a medical point of view, the BIC values calculated by the mclust algorithm are very similar for a wide range of models (Figure 5, bottom left graphic), indicating that from a statistical point of view other models with two, four or more clusters fit almost as well. As shown in the bottom right graphic of Figure 5, in the five-cluster model, mclust splits the collective of the presumably healthy women into two almost identical clusters with reference intervals of 9.8–14.2 and 9.7–14.1 g/dL. This is a typical case of statistical overfitting without substantial gain of information. This finding makes it clear that mclust's proposals, which are based on statistical criteria, are not suitable for making automated estimates of reference limits.

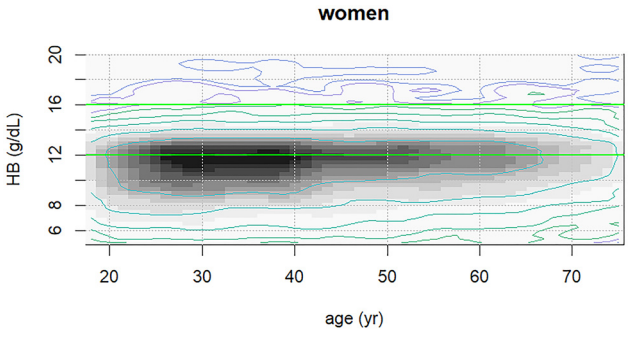

Figure 4: Two-dimensional density plot of hemoglobin values measured in Nigerian women as a function of age. Horizontal green lines represent Caucasian reference limits.

Creatine kinase

In our second experiment, we want to test whether the mixture modelling approach can also be applied to skewed

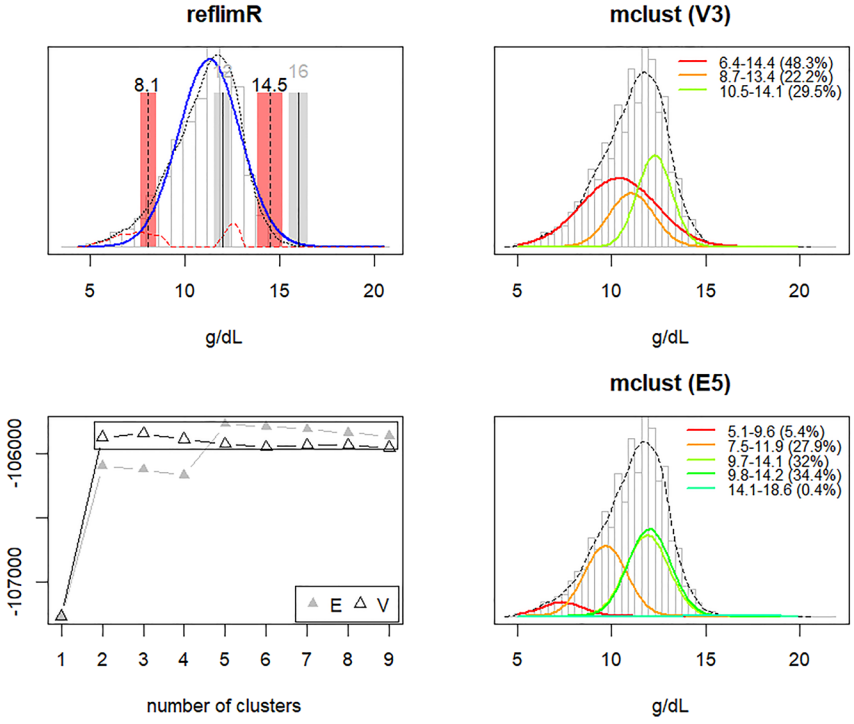

Figure 5: Analysis of hemoglobin values measured with reflimR (upper left) and mclust in Nigerian women. The Mclust function with decomposition of the distribution into three clusters with variable variances (model V3, upper right) or five clusters with equal variance (model E5, bottom right). Very similar Bayesian Information Criterion (BIC) values are obtained for models with less or more clusters (bottom left).

distributions. As described earlier [10], we assume that the right-skewed distribution of creatine kinase (CK) can be approximated to a Gaussian distribution by taking logarithms of the original values.

Similar to hemoglobin in the previous experiment, we first ensure that there is no major age dependence for CK in our study cohort of patients (Figure 6). In contrast to the hemoglobin study (Figure 4), the majority of the CK values (dark region in Figure 6) is within the reference limits of 39–308 U/L specified by the assay manufacturer.

Figure 7 confirms our assumption that the right-skewed distribution of the original CK values (black curve, quartile

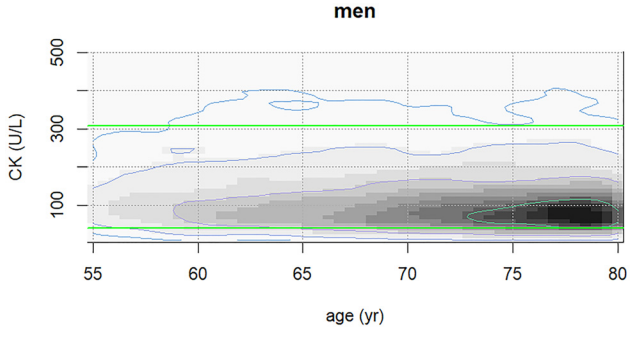

Figure 6: Two-dimensional density plots of real-world creatine kinase (CK) values measured in German hospitalized patients as a function of age. The green lines indicate the target values provided by the assay manufacturer (2.5th and 97.5th percentiles for healthy men).

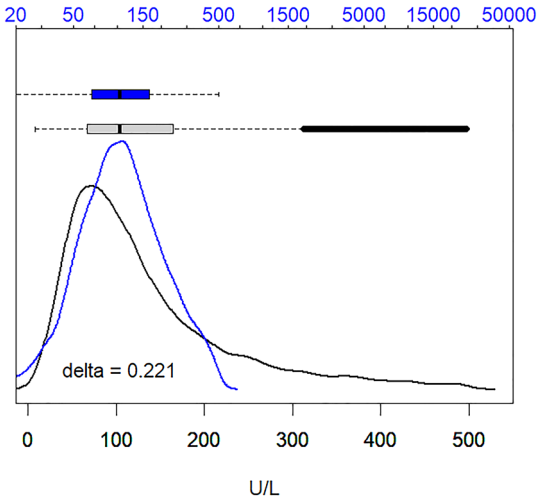

Figure 7: Density curves and boxplots for the original creatine kinase (CK) values (black) and their logarithms (blue) and corresponding skewness delta.

skewness of 0.224) can be modelled lognormal. The log-transformed values show a quite symmetric density curve and boxplot, with a quartile skewness of 0.003. The skewness delta of 0.221 is above the threshold of 0.05 [10] thus indicating that the logarithms of CK values can be modelled by a Gaussian distribution. Consequently, we apply the Mclust function to the logarithms of CK.

Figure 8 shows the results of the reference interval estimation for CK under these assumptions. The reflimR

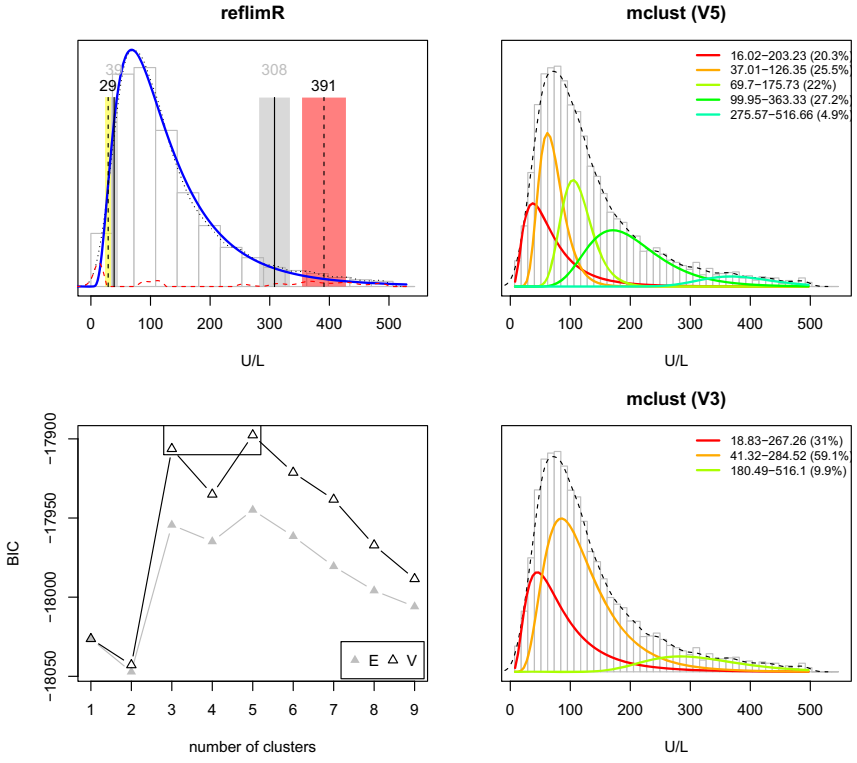

Figure 8: Analysis of real-world creatine kinase (CK) values with reflimR (upper left) and mclust. The Mclust function with decomposition of the distribution into five clusters with variable variances (model V5, upper right) or three clusters with equal variance (model V5, bottom right). The black rectangle outlines models with the highest Bayesian Information Criterion (BIC) values (bottom left).

algorithm assumes 99.5 % unsuspicious CK values and estimates an interval of 29–391 U/L. The corresponding results of the refineR algorithm are 28–289 U/L with 91.3 % unsuspicious values. The reflimR interval deviates markedly from the specifications of the manufacturer. It is too wide and shifted towards the side of the pathological values. The reason is that only 0.5 % of all CK values are assumed to be pathological. The refineR algorithm provides a better estimate of the reference interval because it assumes 8.7 % pathological values.

The mclust algorithm automatically decomposes the overall density curve into five lognormally distributed clusters, three of which fall into the target reference interval. Since none of them confirms the limits specified by the manufacturer, we take a closer look at the model performance of the mclust analysis. With BIC values of –17,905 and –17,898, respectively, the three-cluster and five-cluster models perform almost equally well, whereas two-, four- and six-cluster models come out much less likely (Figure 8, bottom left graphic).

Under medical considerations, the three-cluster model is the most plausible one: it resolves the overall density curve into a non-pathological cohort with a share of about 30 % and a reference interval of 19–265 U/L as well as two pathological cohorts with shares of about 60 and 10 % and upper limits of 285 and 516 U/L, respectively (Figure 8, bottom right graphic). Although the specified reference interval of 39–308 U/L is not met exactly in this experiment, the result is close enough to the target values to confirm our assumption that the Mclust function can be successfully applied to the logarithms of right-skewed distributions.

Discussion

The idea for the Gaussian mixture modelling method presented here came up in 2009, when a team of French veterinarians proposed to split a composite distribution of routinely measured laboratory values into a normal and a pathological subset [13]. Using an expectation maximization (EM) method, they determined the probability of each value to belong to either of the two subsets based on the assumption that the underlying distributions can be modelled as Box-Cox transformed normal distributions with different means and variances.

Although this pioneering approach has been mentioned in some recent reviews [5, 14], it is largely unknown in laboratory medicine, probably because of the complexity of the underlying mathematics and significant limitations in its application to distributions with more than two subgroups [5].

In the present work, we show that the R package mclust simplifies the use of EM methods for Gaussian mixture in a way that specific problems in laboratory medicine can be solved when the aforementioned R packages fail. After

proving the basic feasibility of the approach with simulated Gaussian distributions (Figures 1–4), we demonstrate its practical usefulness in two real-life situations where high proportions of pathological values limit the applicability of the two standard packages reflimR and refineR.

In the two examples shown in Figures 5 and 8, standard methods estimate reference limits that are too wide and shifted to the side of the pathological values. The reason is that they include too many pathological values in the presumably normal collective. In both cases, the mclust analysis suggests only 30 % of all values to belong to the normal collective. This low proportion overtaxes all indirect methods for reference interval estimation, which rely on the assumption that the non-pathological values represent the majority of all measurements [3–5, 8, 9].

In the haemoglobin experiment, mclust supports the hypothesis that the lower limit set by the aforementioned indirect methods is too low. This is due to a mixture of strongly overlapping normal distributions that cannot be decomposed. Many African countries suffer from a high proportion of anemic patients [15], which is also true for Lagos, Nigeria [16], where this study was conducted. For that region, Akinbami et al. report a hemoglobin reference interval as low as 9.9–14.7 g/dL [16], which is in good agreement with one of the fractions identified by the Mclust function in the present study (10.5–14.1 g/dL). Once again, we would like to emphasize that mclust is not intended for automatically estimating plausible reference interval. Instead, it offers different solutions for Gaussian mixture deconvolution based on high BIC values. From these solutions, a medically plausible explanation can be derived as to why standard methods struggle with strongly overlapping distributions. The decision of which model to choose is the responsibility of the experienced medical professional, not the software.

There is remarkable disagreement in the literature about the exact reference limits for hemoglobin, even when these were determined on relatively homogeneous Caucasian populations [17]. It is therefore even more important not to uncritically transfer the specifications of Western IVD manufacturers to other continents [18]. Published reports are scarce, and the results are divergent [16, 18–20]. With respect to hemoglobin in Africa, published intervals range from 8.1 – 14.2 g/dL in rural Western Kenya [19] to 12.5–17.6 g/dL in Asmara, Eritrea. This enormous difference may be partly due to genetic and nutritional factors in Africa [12] and partly to the difference in altitude, as Asmara is located over 2,300 m above sea level. These widely differing conditions clearly indicate that reference intervals generated in Caucasians should by no means be regarded as standard for the rest of the world.

In our second example, we demonstrate on the one hand that the mclust can also be applied to skewed distributions and on the other hand that it can also provide useful results when indirect methods for reference interval estimation must be applied to hospitalized patients with a high proportion of pathological values. We chose CK measurements in a center with a high prevalence of cardiac diseases. Here, mclust succeeds in splitting the patients into cases with presumably normal, borderline and distinctly pathological values.

Similar conditions will be encountered in many routine situations in everyday medical practice, for example when estimating reference intervals for point-of-care devices, on which mainly pathological values are monitored, or in special departments such as endocrinology where patients with specific medical issues are examined.

Conclusions

The standard R packages reflimR and refineR are generally sufficient for estimating reference intervals from the laboratory values of mixed patient collectives, as long as the proportion of pathological values does not exceed 25 % on either side of the distribution [8, 9]. However, there are situations in which this requirement cannot be met. This is where machine learning with Gaussian Mixture Modeling [6] can make a valuable contribution to pointing out any inhomogeneities in the data that are not recognized or cannot be resolved by the standard methods. In such cases, mclust can estimate the proportions of the various fractions and provide at least a rough idea of the correct reference limits.

The BIC analysis of mclust automatically provides the model with the best fit to the overall distribution of measurements. But in many cases, there are several solutions with similarly high BIC values. Therefore, it is highly recommended to consider the absolute BIC values with statistical and medical expertise and select the model that best addresses the practical issue.

In particular, mclust often provides similarly good fits for distribution mixtures with either equal or different variance (models E and V). From a medical point of view, “diseased” vs. “healthy” subpopulations would likely not have the same variance as they stem from distinct states of pathophysiology. Therefore, in case of doubt, the decision should be made in favor of a model with variable variance, even if this has larger overlaps between the fractions.

This brief discussion shows that Gaussian mixture models are a valuable addition to standard methods for RI

verification and should be further explored and improved in the future.

Research ethics: Research involving human subjects complied with all relevant national regulations, institutional policies and is in accordance with the tenets of the Helsinki Declaration and has been approved by the National Health Research Ethics Committee (NHREC), in Nigeria (19/12/2008a; Health Research Committee assigned number: ADM/DSCST/HREC/APP/6420) and an ethics vote in respect of the use of routine data for quality control is given by the Technical University Munich (10/21 S-KH).

Informed consent: Not applicable.

Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: The authors state no conflict of interest.

Research funding: None declared.

Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Horowitz GL, Altaie S, Boyd JC, Ceriotti F, Garg G, Horn P, et al. EP28-A3C: Defining, establishing, and verifying reference intervals in the clinical laboratory; Approved Guideline - Third Edition. Vol. 28. No. 30. 3rd ed. Wayne: Clinical and Laboratory Standards Institute; 2010.
2. Ozarda Y, Higgins V, Adeli K. Verification of reference intervals in routine clinical laboratories: practical challenges and recommendations. Clin Chem Lab Med 2018;57:30–7.
3. Jones G, Haeckel R, Loh T, Sikaris K, Streichert T, Katayev A, et al. Indirect methods for reference interval determination: review and recommendations. Clin Chem Lab Med 2018;57:20–9.
4. Ozarda Y, Ichihara K, Jones G, Streichert T, Ahmadian R. Comparison of reference intervals derived by direct and indirect methods based on compatible datasets obtained in Turkey. Clin Chim Acta 2021;520:95.
5. Haeckel R, Wosniok W, Streichert T. Review of potentials and limitations of indirect approaches for estimating reference limits /intervals of quantitative procedures in laboratory medicine. J Lab Med 2021;45:35–53.
6. Scrucca L, Fraley C, Murphy TB, Raftery AE. Model-Based Clustering, Classification, and Density Estimation Using mclust in R. Chapman and Hall/CRC; 2023. Available from: https://cran.r-project.org/web/packages/mclust.
7. Arzideh F, Özcürümez M, Albers E, Haeckel R, Streichert T. Indirect estimation of reference intervals using first or last results and results from patients without repeated measurements. J Lab Med 2020;45:103–9.
8. Hoffmann G, Klawitter S, Trulson I, Adler J, Holdenrieder S, Klawonn F. A novel tool for the rapid and transparent verification of reference intervals in clinical laboratories. J Clin Med 2024;13:4397.
9. Ammer T, Schützenmeister A, Prokosch HU, Rauh M, Rank C, Zierk J. refineR: a novel algorithm for reference interval estimation from real-world data. Sci Rep 2021;11:16023.
10. Klawonn F, Hoffmann G, Orth M. Quantitative laboratory results: normal or lognormal distribution. J Lab Med 2020;44:143–50.
11. Nebe T, Bentzen F, Bruegel M, Fiedler M, Gutensohn K, Heimpel H, et al. Multicentric dertermination of reference ranges for automated blood counts. J Lab Med 2011;35:3–28.
12. WHO. The global prevalence of anemia in 2011. Geneva: World Health Organization; 2015. Available from: https://www.who.int/publications/i/item/9789241564960.
13. Concordet D, Geffré A, Braun J, Trumel C. A new approach for the determination of reference intervals from hospital-based data. Clin Chim Acta 2009;405:43–8.
14. Holmes D, Buhr K. Widespread incorrect implementation of the Hoffmann method, the correct approach, and modern alternatives. Am J Clin Pathol 2019;151:328–36.
15. Nti J, Afagbedzi S, Vroom F, Ibrahim N, Guure C. Variations and dterminants of anemia among reproductive age women in five sub-Saharan Africa countries. BioMed Res Int 2021;2021:9957160.
16. Akinbami A, Wright K, Dada A, Uche E, Dada A, Ouwah A, et al. Population-based full blood count reference intervals in Lagos, Nigeria. Ann Trop Pathol 2023;14:16–22.
17. Haeckel R, Arzideh F. Common reference intervals of blood counts. J Lab Med 2011;35:217–8.
18. Siraj N, Anwar M, Mehari Y, Russom S, Kahsay S, Frezghi H, et al. Establishment of hematological reference intervals for healthy adults in Asmara. BMC Res Notes 2018;11:55.
19. Zeh C, Amornkul P, Inzaule S, Ondoa P, Oyaro B, Mwaengo D, et al. Population-based biochemistry, immunologic ad hematological reference values for adolescents and young adults in a rural population in Western Keny. PLoS One 2011;6:e21040.
20. Saathoff E, Schneider P, Kleinfeldt V, Geis S, Haule D, Maboko L, et al. Laboratory reference values for healthy adults from southern Tanzania. Trop Med Int Health 2008;13:612–25.