Article

Univariate Analysis of Short-Chain Fatty Acids Related to Sudden Infant Death Syndrome

Carlos E. Galván-Tejada [1,†] , **Karen E. Villagrana-Bañuelos** [1,†] , **Laura A. Zanella-Calzada** [2,†] ,
Arturo Moreno-Báez [1] , **Huizilopoztli Luna-García** [1] , **Jose M. Celaya-Padilla** [1] ,
Jorge I. Galván-Tejada [1] and **Hamurabi Gamboa-Rosales** [1,*]

[1] Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Jardín Juarez 147, Centro, Zacatecas 98000, Mexico; ericgalvan@uaz.edu.mx (C.E.G.-T.); kvillagrana@uaz.edu.mx (K.E.V.-B.); morenob20@uaz.edu.mx (A.M.-B.); hlugar@uaz.edu.mx (H.L.-G.); jose.celaya@uaz.edu.mx (J.M.C.-P.); gatejo@uaz.edu.mx (J.I.G.-T.)

[2] LORIA (INRIA, CNRS), Campus Scientifique BP 239, Université de Lorraine, 54506 Nancy, France; laura.zanella-calzada@univ-lorraine.fr

* Correspondence: hamurabigr@uaz.edu.mx; Tel.: +52-4921216787

† These authors contributed equally to this work.

Abstract: Sudden infant death syndrome (SIDS) is defined as the death of a child under one year of age, during sleep, without apparent cause, after exhaustive investigation, so it is a diagnosis of exclusion. SIDS is the principal cause of death in industrialized countries. Inborn errors of metabolism (IEM) have been related to SIDS. These errors are a group of conditions characterized by the accumulation of toxic substances usually produced by an enzyme defect and there are thousands of them and included are the disorders of the β-oxidation cycle, similarly to what can affect the metabolism of different types of fatty acid chain (within these, short chain fatty acids (SCFAs)). In this work, an analysis of postmortem SCFAs profiles of children who died due to SIDS is proposed. Initially, a set of features containing SCFAs information, obtained from the NIH Common Fund's National Metabolomics Data Repository (NMDR) is submitted to an univariate analysis, developing a model based on the relationship between each feature and the binary output (death due to SIDS or not), obtaining 11 univariate models. Then, each model is validated, calculating their receiver operating characteristic curve (ROC curve) and area under the ROC curve (AUC) value. For those features whose models presented an AUC value higher than 0.650, a new multivariate model is constructed, in order to validate its behavior in comparison to the univariate models. In addition, a comparison between this multivariate model and a model developed based on the whole set of features is finally performed. From the results, it can be observed that each SCFA which comprises of the SFCAs profile, has a relationship with SIDS and could help in risk identification.

Keywords: sudden infant death syndrome; short-chain fatty acids; univariate analysis; generalized linear model

1. Introduction

Sudden infant death syndrome (SIDS) is defined as the death of a child under one year of age, during sleep, without apparent cause. Several investigations have been carried out and the concept of SIDS has changed in order to discover the specific reason that causes death, thus it can be defined as the sudden death of an infant under one year of age, while asleep, which remains unexplained after a thorough investigation, including performance of a complete autopsy, examination of the death scene, and review of clinical history [1–4].

According to the Centers for Disease Control and Prevention (CDC) in the United States there are 3600 sudden unexpected infant deaths [5]. In Mexico there are around 350 child deaths per year, corresponding to approximately one out of every two thousand born, during the first year of life [6]. Unfortunately in Mexico, as in other countries, the statistics on the frequency of SIDS are not real, either due to under-registration or misdiagnosis [7].

SIDS is the main cause of death in previously healthy babies [8], as of now there is no consensus on the reasons that causes it [9] which results in uncertainty when not being able to identify which babies are at risk of suffering it, and therefore of suffering a lethal episode, also known as episodes of apparent life threatening events (ALTE, "an episode that is frightening to the observer and that is characterized by some combination of apnea (central or occasionally obstructive), color change (usually cyanotic or pallid but occasionally erythematous or plethoric") [10].

There are data on SIDS, but they provide no explanation for the etiology and pathogenesis of sudden infant death. Sudden infant death is considered "multifactorial" [11] and this concept, however, helps as little as the term "idiopathic", which is used in many other diseases to explain that something does not have an explication yet. As our knowledge of SIDS grows, it will be possible to recognize and classify pathological conditions behind the sudden death of a child [12].

Several investigations have been carried out from all possible areas, such as engineering [13,14], epidemiology [15], pathophysiology [16], pathology [2], pediatrics [17], neurology [18], among others in the search for solutions to it. Among the associated conditions, inborn errors of metabolism (IEM) stand out [19,20] .

IEM are a group of conditions characterized by the accumulation of toxic substances usually produced by an enzyme defect [21]. It can also be defined as a diverse heterogeneous group of disorders with protean clinical manifestations presented mainly in the pediatric population [22]. Disorders of the β-oxidation cycle are included in the group of diseases that include IEM: Acetyl CoA is generated from fatty acids through repeated cycles of β-oxidation. Groups of four specific enzymes are required for different chain lengths (very long chain, long chain, medium chain, and short chain) [21]. Thus one of the pathologies that could explain SIDS are the enzymatic defects to metabolize short chain fatty acids (SCFAs), these being comprised of 1–6 carbon based anions, which are produced during bacterial fermentation, of which, acetate (C2), propionate (C3), and butyrate (C4) are the most abundantly produced. SCFAs are produced naturally within the colon by fermentation of carbohydrates, both dietary and endogenous, and proteins that are accessible to the microbiota [23].

There are some investigations that have related the IEM with SIDS or ALTE [20,24], Works such as those of Villoria et al. [25], mention that short-chain fatty acid problems are considered among the possible causes of this syndrome, however the statistics between the relationship are not established due to methodological problems.

IEM are currently considered treatable pathologies if they are diagnosed promptly [26]. In the pediatric age from the first hours of life they can show symptoms and signs similar to other diseases, or even have no symptoms. which makes them difficult to diagnose, this leads to sequelae such as malnutrition, seizures, mental retardation, apnea [24], and ALTE, which could also compromise consciousness (lethargy and drowsiness, which leads to deep coma and death) [24].

The incidence of IEM as a cause of SIDS or ALTE remains unknown (or, at least, not well known) [24], however if proper investigations are not carried out, these cases often meet the criteria for SIDS [27]. There are controversies regarding the proportion of SIDS attributable to fatty acid oxidation defects. It is mentioned that there are no reliable records of SIDS, the investigations could have biased data such as differences in the selection of cases, and the type of methodology, among others [27]. As previously mentioned, when diagnosed these disorders can be treated in a timely manner.

SIDS, as mentioned before, is a diagnosis of exclusion, although it means that the reasons why this occurs have not been discovered yet. With the advent of advances in science, opportunities are opened to observe, analyze, and find relationships among the possible risk factors.

A proposal that has increased the analysis of relationships between variables and SIDS is the application of artificial intelligence techniques, such as machine learning (ML), looking for the development of computed-aided diagnosis (CAD) tools, and seeking to counter the points mentioned above. Kononenko, I. [28] mentioned how artificial intelligence, which is part of computing, tries to make computers smarter based on learning. Most of the current research agrees that it is not smart but that it learns. That is why to create this type of learning model, based on the capacity of what was previously learned, it is necessary that the data provided to the computer be reliable, as complete as possible for an adequate characterization of the disease and so to predict and diagnose.

In this sense, CADs have demonstrated to be useful in many fields of human medicine to aid in diagnostics [29–32], but it is important to highlight that their role in pediatrics; Kokol et al. [33] has detected the relevance that artificial intelligence has had in the pediatric area, since 2013 and beyond that has focused on the application of ML in diseases like schizophrenia, pneumonia, asthma, abnormality, and epilepsy. No new ML approaches emerged, however the assessment of accuracy with the introduction of new metrics has become important and the research focus has shifted from classification to predictive models. For example, Mitchell et al. [33] developed a predictive model in children with acute lymphoblastic leukemia. It consisted of a predictive model for identifying the increased risk for thromboembolism. It would be beneficial in targeting interventional studies to high-risk groups. It was evaluated in 456 children and then validated in 339 patients and the model's specificity was 96.2% and its sensitivity was 63.2%, on the basis of high specificity, the model may identify children with leukemia at risk of thromboembolism. Mueller et al. [34] used a dataset of 486 mechanically ventilated premature infants to develop predictive models and determine whether machine learning methods can predict the extubation outcome in premature infants as well as clinicians using machine learning algorithms such as artificial neural networks (ANN), support vector machine (SVM), naive Bayesian classifier (NBC), powered decision trees (BDT), and multivariate logistic regression (MLR). Finding that for some models (ANN, MLR, and NBC) in the IEM study, the ML techniques also had satisfactory results (area under the receiver operating characteristic curve (AUC): 0.63–0.76). There are works that use classification algorithm methodologies with random forest (RF) and include metabolic, socioeconomic, demographic characteristics, among others, to help with the prognosis, such as the work of Wandhwani et al. [35] who applied the above to know the prognosis in the liver transplants of pediatric patients.

Therefore, in this work the analysis of postmortem SCFA values of babies with and without SIDS, based on the development of univariate models through a generalized linear models approach (GLM), is proposed as the main objective. Based on an univariate statistical analysis and a multivariate comparison, the intention is to contribute to a future assisted diagnosis to support in the diagnosis of the risk of suffering ALTE without reaching SIDS.

This paper is organized as follows. Section 2 presents a detailed description of the SCFA profile of infants data set used, as well as the methods applied for the univariated analysis to find relationships. Section 3 presents the experiments performed using the eight short fatty acids as well as the results of the logistic regression and the evaluation of each with the receiver operating characteristic curve (ROC curve) and AUC. In Section 4, a discussion about the results of the relation of SFCA and SIDS is described and presents the conclusions of the work.

2. Materials and Methods

In this study, a public database available from the NIH Common Fund's National Metabolomics Data Repository was used. This study is summarized as an *Analysis of SCFA profile in infants dying of SIDS compared to infants dying of controls* from the University of Michigan, Biomedical Research Core Facilities [36]. The criteria to select this dataset were:

1.   Case-control study;
2.   Short-chain fatty acids profile for cases and controls;
3.   SIDS related study.

Given these constraints, only one public dataset was found. This dataset is described in detail below.

2.1. Data Description

The dataset denominated "SCFA profile in babies dying from SIDS" [36], containing information referred to 18 patients, 5 controls, and 13 patients diagnosed with SIDS. This information includes the values obtained for the features presented in Table 1 which were calculated postmortem through cold extraction measured by electron ionization-gas chromatography mass spectrometry (EI GC-MS) without derivatization. The unit of these values is uM. For more details on the process that was used to prepare the sample, the description in the original project can be consulted [36].

Table 1. Features contained in the "short chain fatty acids (SCFAs) profile in babies dying from sudden infant death syndrome (SIDS)" dataset.

	Feature
1	Postmortem interval PMI hours
2	Gestational age weeks
3	Postnatal age weeks
4	Isovaleric acid
5	Octanoic acid
6	Propionic acid
7	Isobutyric acid
8	Butyric acid
9	Hexanoic acid
10	Valeric acid
11	Acetic acid

2.2. Data Availability

This data is available at the NIH Common Fund's National Metabolomics Data Repository (NMDR) website, the Metabolomics Workbench, https://www.metabolomicsworkbench.org where it has been assigned Project ID PR000512. The data can be accessed directly via the Project DOI: 10.21228/M8GQ27. This work is supported by NIH grant U2C-DK119886 [36].

2.3. Data Preprocessing

Data pre-processing consisted of forming a tidy data, which included each of the short chain fatty acids and the postmortem interval. In addition, the gestational age and postnatal age information was analyzed, identifying that these data did not agree with reality, since the gestational age at birth ideally covers a maximum of 42 weeks of gestation, so that when these weeks are exceeded, it puts the life of the binomial mother-son at risk [37]. Therefore, a possible error was identified in the labeling of said data, so the original data for said ages, provided by NIH Common Fund, were inverted, changing the values referring to the gestational age for those referring to postnatal age and the opposite.

2.4. Generalized Linear Model

A GLM represents a generalization of a linear model through expanding its scope to nonlinear data that can be transformed into a linear form using suitable transformations. GLMs allow one to handle the limitation that linear regression presents, which is the assumption of linear relationships between the input and output. The nonlinear relationship between the input and output is converted into linear by adding a step of transforming part of the data (the input or the output) into another domain. This step is known as the basis function. One of the basis functions that is widely used is logistic regression, which uses a logistic function to transform the nonlinearity into linear. In the logistic function, the output is also mapped between a range of [0, 1], being equivalent to a probability density function.

In logistic regression, an exponential functional to the linear regression output is added, $y_i \in \Re$, to constrain it into $y_i \in [0, 1]$. The relationship between the input and predicted output can be calculated with Equation (1).

$$\hat{y}_i = \sigma(\sum_{j=1}^{n} x_{i,j}.w_j + w_0).$$ (1)

As the output obtained by logistic regression presents a symmetrical distribution between $[-\infty, \infty]$, it is better suited for classification problems [38].

2.5. Validation

In this study, there are only two possible outputs for the classifier, "positive" and "negative". These possible outputs can be represented within a confusion matrix, which allows one to observe the differences between the positive and negative predicted by the classifier, as well as the correct class of each case.

This matrix allows an observation of true positives (TP), true negatives (TN), as well as the false positives (FP) and false negatives (FN). Based on this, different metrics can be estimated to help measure the performance of the model under different criteria. Specifically, there are two metrics that can be calculated. Firstly, sensitivity can be calculated by Equation (2):

$$Sensitivity = \frac{TP}{TP + FN},$$ (2)

which describes the proportion of true positives, i.e., the probability that a case of SIDS will be correctly classified. And secondly, specificity [39], which can be calculated by Equation (3),

$$Specificity = \frac{TN}{FN + TP},$$ (3)

which describes the proportion of true negatives, that is, the probability that a case of non-SIDS will be correctly classified.

Therefore, to evaluate the results of the proposed analysis, an approach was carried out using the ROC curve. The ROC curve is a well known technique based on the metrics mentioned before [40], used to visualize the performance of a classifier, complemented further by the area under the ROC curve (AUC), which is dependent on a decision point (threshold). The AUC value represents the probability that a random positive sample is correctly identified and it represents a desirable measure since it is scale-invariant (measuring the ranking of the predictions instead of their absolute values) and its classification-threshold invariant (measuring the prediction's quality irrespective of the threshold selected).

3. Experiments and Results

This section presents the experiments performed for the development of this work as well as the results obtained.

Initially, the set of features contained in the dataset were submitted to an univariate analysis, developing a model based on the relationship between each feature and the outcome, obtaining 11 univariate models. Then, each model was validated calculating their ROC curve and AUC value. For those features whose models presented an AUC value higher than 0.650, a new multivariate model was constructed, in order to validate its behavior in comparison to the univariate models. In addition, a comparison between this multivariate model and a model developed based on the whole set of features is finally performed.

Table 2 presents the 11 features used for this work together with the AUC values they obtained through the univariate analysis performed, based on a GLM approach. Features are listed in ascending order according to the AUC value calculated for each.

Table 2. AUC values obtained by each feature through the univariate analysis.

Feature	AUC Value
Isovaleric acid	0.508
Octanoic acid	0.538
Gestational age weeks	0.592
Postmortem interval PMI hours	0.600
Propionic acid	0.646
Isobutyric acid	0.662
Postnatal age weeks	0.662
Butyric acid	0.723
Hexanoic acid	0.769
Valeric acid	0.815
Acetic acid	0.846

Figure 1 presents the ROC curves corresponding to the univariate models that reached an AUC value higher than 0.650. In (a) it is shown that the ROC curve belonging to the model developed for the isubutyric acid feature, in (b) for the postnatal age weeks feature, in (c) for the butyric acid feature, in (d) for hexanoic acid, in (e) for valeric acid feature and, and in (F) for acetic acid.
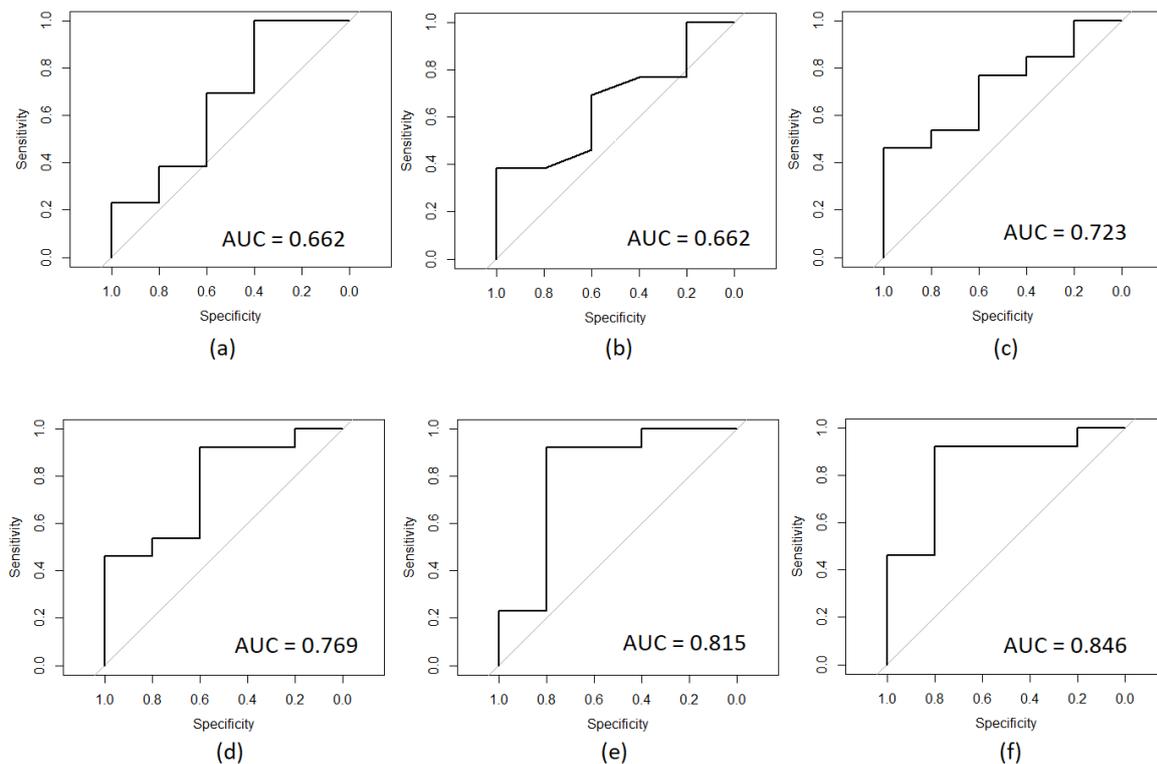
Figure 1. ROC curves obtained through the univariate models of the features that presented an AUC value higher than 0.650: (**a**) Isubutyric acid, (**b**) postnatal age weeks (**c**) butyric acid, (**d**) hexanoic acid, (**e**) valeric acid, and (**f**) acetic acid.

Figure 2 presents the ROC curve obtained for the model based on the total set of features in (a), while (b) presents the ROC curve obtained for the model based on the set of features that presented an AUC value higher than 0.650.
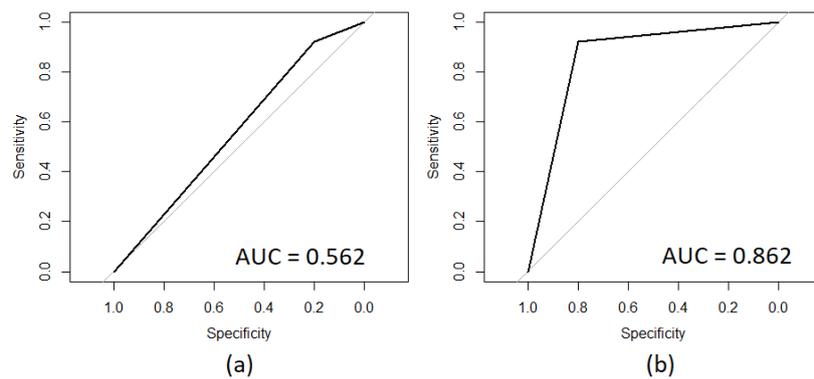
Figure 2. ROC curves obtained through (**a**) the model containing the complete set of features and (**b**) the model containing the set of features presenting an AUC value higher than 0.650.

From the results, it can be observed that each SCFA which comprise the SCFAs profile has a relationship with SIDS, according to the validation metric used, AUC, presented in Table 2. However, only the models of those features that presented an AUC value higher than 0.650 were taken into account for a subsequent multivariate evaluation, since the probability that they correctly identify a positive patient is higher than 0.65%, being statistically significant. For these features, the ROC curve is shown in Figure 1, where it can be observed that not only did the series of fatty acids, isobutyric, butyric, hexanoic, valeric, and acetic, present a significant value for the identification of positive patients, but also for the postnatal age in weeks. Regarding these SCFAs, the valeric acid and acetic acid were those that obtained higher AUC values, presenting 0.815 and 0.846, respectively, showing that based on the information contained in these features, the probability that a positive patient was correctly identified is higher than 81%. On the other hand, the isobutyric acid was the one that obtained a lower AUC, presenting a value of 0.662, meaning a patient who was identified as positive has a 60% probability of having a correct classification. It is remarkable that the feature containing information about the postnatal age obtained an AUC value equal to the one obtained by the feature of the isobutyric acid, meaning that both are presenting the same contribution in the correct identification of positive patients, according to this value however, according to the ROC curves, it can be seen that the behavior of these is evidently different, which represents that the contribution of both features is not applied uniformly. That is to say that while the model with the isobutyric acid might be correctly identifying a set of patients, the model of the postnatal age in weeks might be correctly identifying a different set.

The AUC descriptor was selected because it is presented in the bioinformatics and biomedical literature [41–45] as a tool to describe the behavior of the phenomenon, giving a metric that, if it is greater than 0.5 it is considered useful for describing the phenomenon, in this case the relationship of short-chain fatty acids and SIDS.

Therefore, based on the results it is notable that the behavior of a model developed from the full set of features does not generate any contribution for the identification of positives. However, this does not mean that the features by themselves do not contain information that is representative for the identification of positives, but that modeled together they do not achieve significant generalization. Therefore, if an analysis is carried out univariately to identify the contribution of each feature, in order to know which ones are representative for the correct identification of positives, as was done in this work, it is possible to rule out those that do not generate a contribution and keep only those that do. In this way, the number of features necessary for the identification of the target is reduced, eliminating unnecessary, redundant, or non-contributing information, and a generalized model is developed, allowing one to identify SIDS with a significant probability of correctness.

4. Discussion and Conclusions

In this work, it was identified that a multivariate model, as presented in other bioinformatic and biomedical approaches research [46–48], presented better results compared to an univariate model. However, the univariate model was necessary to identify the contribution of each characteristic and conform the multivariate model. It was shown that the characteristics of babies who died from SIDS, using values of short chain fatty acids, were able to classify cases of death for SIDS. Therefore, these results show that, in addition to the SCFAs that better relate to SIDS, the age of the infant in weeks contributes in the identification of positives.

It stands out, that if a patient is labeled as SIDS, after of exhaustive investigation according to the definition in the literature, metabolic diseases were ruled out. The results were interesting, due to low probability rate of metabolic diseases [49], even more, if the subject was subjected to medical research, as the criteria to classify it as death from SIDS mark it [50]. Nevertheless, there are discrepancies between countries regarding the protocols for investigating cases of death from SIDS [51].

The relationship between IEM and potential death from SIDS is well known and has been described in the literature for three decades [25]. However, it is important to point out the high discrepancies between countries for neonatal screening of metabolic diseases and as there are thousands of them [25], to examine each patient for each metabolic disease. In Europe, national programs, the screening of metabolic diseases vary by region ranging from two up to 30. In addition regions of America, Asia, the Middle East, and North Africa [25], present similar situations. The metabolic disorders that are investigated in most countries include: Phenylketonuria, congenital hypothyroidism, congenital adrenal hyperplasia, cystic fibrosis, galactosemia, biotinidase deficiency, and acyl-CoA deficiency medium chain dehydrogenase [52]. As can be seen, programs in search of diseases related to defects in SCFA metabolism are not routinely included in most countries. In addition, due to the nature of these disorders and the complexity of both the type of substances and chemical steps in a certain metabolic pathway, they are not suspected and are hardly diagnosed [53]. It is relevant that pathologies such as medium chain fatty acid disorders are analyzed in less than fifty percent in regions of Europe and America [25], being one of the main disorders associated with SIDS, which allows one to deduce that a small percentage of research is carried out for short chain disorders. It would be convenient to take these variables into consideration in subsequent studies. Another point to consider is that a part of physiological regulation involves metabolites such as SCFAs, which is accessed through the intestinal epithelium interacting with the cells. These acids not only serve as energy for the intestinal microbiota, but also for the intestinal epithelial cells, having functions of regulators in the physiology and immunity of the host, resulting in beneficial metabolites with anti-inflammatory capacity [54]. Furthermore, breastfeeding could be associated with the concentrations of short chain fatty acids, either by fetal maternal transfer, or by the type of bacterial colonization of the colon, which is widely recognized as an immune benefit of exclusive breastfeeding, which could contribute to the risk of suffering SIDS [55]. Together with this research, this supports that the disorders related to SCFAs should be investigated, since they have a fundamental role in immunity and are able to characterize SIDS through the multivariate model.

Nevertheless, there are several issues that must be considered in this study:

- This case-control study was performed using a dataset from the University of Michigan, in the Boston children's Hospital. From the literature, we are aware that NBS differ methodologically and in the disorders screened worldwide, meaning that this study can not be generalized onto the worldwide population. Several factors must be considered when the results are generalized, among them the profiles of SCFAs that were obtained to perform the data-set, which were postmortem. Thus, it is unknown if the found values fluctuate as the postmortem interval increases, and therefore whether they differ or not from a living subject. It is also recognized that race could also influence these values, since it has been described in various investigations that the prevalence is higher in certain race groups as opposed to the white race. In addition

to racial characteristics, environmental influence, population lifestyles, access to health services (prenatal care, childbirth care, well-child care), socioeconomic status, among many other factors that vary according to the population, the territory must also be considered, which can directly and indirectly affect the appearance of SIDS;

- As described in Section 2, this study is comprised of 18 subjects, which is prone to overfitting even when a blind-test approach is performed to validate experimentation. As this is a small number of cases, valuable features for case identification could be excluded. One of them is that most of the study subjects were close to one year of age, so there could be a risk of overfitting and not identifying younger patients, since a higher range of risk has been identified in the literature, which is between the ages of 2 and 8 months, data that varies according to the authors, this age could mark significant differences in SCFAs levels according to weeks of life;

- As presented in Section 3, the SCFAs profile gives us insights to possible SIDS complications, however, we are aware that SIDS is a complex, multifactorial disorder, which can be influenced by other risk factors, it being a disease described as a syndrome and remembering that the concept refers to a set of signs and symptoms that characterize a disease, so the absence or presence of a particular sign or symptom is not decisive for suffering it. However, since several symptoms or signs are present in a patient, they have become relevant for their study. Until now, the most significant factors that describe this pathology have not been identified. Thus, long lists grouped into genetic or inheritance factors, maternal factors, environmental factors, and newborn factors can be found, just to mention some. It has been shown that each of them could intervene in the presentation of SIDS;

- Therefore, other clinical data from patients is not available in this study and can influence the results. Among which stands out the way of obtaining the product of conception, that is, by delivery or cesarean, which when compared will have different types of bacterial colonization, and remember that the production of SCFAs depends largely on the fermentation of food by part of the bacteria in the colon. This is also influenced by the type of feeding, whether exclusively breastfeeding, supplementary, or combinations of both. In the same sequence of ideas, the age of ablation and weaning have an important role in variations in the gut microbiota among newborns.

Its possible to conclude that the SFCAs profile allows the identification, in a quantitative way, of the risk of SIDS. The risk identification could be done with acetic acid, followed by valeric acid. In addition, a particular SFCAs profile comprised by isobutyric, butyric, hexanoic, valeric, and acetic acids, complemented by the postnatal age in weeks, could improve risk identification. The delimitation of which risk factors are most significant for a possible appearance of SIDS is very relevant, and the foregoing opens opportunities for the identification of risk factors with greater statistical significance and, therefore, useful for creating, in the future, models with greater complexity that involve a greater number of variables, so as to be able to predict with greater certainty the risk of suffering ALTE or SIDS. Research in this sense will become even more relevant as it is useful information for the health sector and society in general, which will make it possible to strengthen prevention and create timely diagnosis strategies that contribute to reducing the mortality rate.

References

1.　Bajanowski, T.; Brinkmann, B.; Vennemann, M. The San Diego definition of SIDS: Practical application and comparison with the GeSID classification. *Int. J. Leg. Med.* **2006**, *120*, 331–336. [CrossRef]

2.　Shipstone, R.; Young, J.T.J.; Byard, R. An evaluation of pathologists' application of the diagnostic criteria from the San Diego definition of SIDS and unclassified sudden infant death. *Int. J. Leg. Med.* **2019**. [CrossRef]

3.　Oyarzún, M.; Brockmann Veloso, P. *Pediatric Respiratory Diseases. Sudden Infant Death Syndrome*, 1st ed.; Pediatric Respiratory Diseases; Springer: Cham, Switzerland, 2020; pp. 495–500.

4.　Krous, H. Sudden Infant Death Syndrome and Unclassified Sudden Infant Deaths: A Definitional and Diagnostic Approach. *Pediatrics* **2004**, *114*, 234–238. [CrossRef]

5.　CDC. Data and Statistics for SIDS and SUID|CDC, 2019. Available online: https://www.cdc.gov/sids/data.htm (accessed on 25 March 2020).

6.　Informativo, P. Uno de Cada dos Mil Bebés Fallece Por la Llamada Muerte de Cuna. Technical Report. 2016. Available online: https://planoinformativo.com/449892/uno-de-cada-dos-mil-bebes-fallece-por-la-llamada-imuerte-de-cunai-slp (accessed on 1 December 2019).

7.　Flores-Huerta, S. Síndrome de muerte súbita del lactante. Prevención en la práctica hospitalaria. *Rev. Med. Inst. Mex. Seguro Soc.* **2006**, *44*, 511–518.

8.　Bryant, V.; Sebire, N. *Natural Diseases Causing Sudden Death in Infancy and Early Childhood, Sudden Infant and Early Childhood Death: The Past, the Present and the Future*, 1st ed.; University of Adelaide Press: Adelaide, Australia, 2018; pp. 539–588.

9.　Goldstein, R.; Blair, P.; Sens, M.; Shapiro-Mendoza, C.; Krous, H.; Rognum, T.; Moon, R. Inconsistent classification of unexplained sudden deaths in infants and children hinders surveillance, prevention and research: Recommendations from The 3rd International Congress on Sudden Infant and Child Death. *Forensic Sci. Med. Pathol.* **2019**. *15*, 622–628. [CrossRef]

10.　Fu, L.; Moon, R. Apparent Life-threatening Events (ALTEs) and the Role of Home Monitors. *Pediatr. Rev.* **2007**, *28*, 203–208. [CrossRef] [PubMed]

11.　Carlin, R.; Moon, R. Risk Factors, Protective Factors, and Current Recommendations to Reduce Sudden Infant Death Syndrome: A Review. *JAMA Pediatr.* **2017**, *171*, 175–180. [CrossRef] [PubMed]

12.　Zubov, L.; Bolganov, Y.; Valkov, A. Sindrom Vnezapnoy Detskoy Smerti. *Med. Neotdozhiykh Sostoyaniy* **2007**, *2*, 114–124.

13.　Liu, X.; Takeuchi, K.; Ogunfunmi, T.; Mathapathi, S. Video-based IoT baby monitor for SIDS prevention. In Proceedings of the 2017 IEEE Global Humanitarian Technology Conference (GHTC), San Jose, CA, USA, 19–22 October 2017; pp. 1–7.

14.　Jayatilleka, I.; Halgamuge, M.N. Chapter 1—Internet of Things in healthcare: Smart devices, sensors, and systems related to diseases and health conditions. In *Real-Time Data Analytics for Large Scale Sensor Data*; Advances in Ubiquitous Sensing Applications for Healthcare; Das, H., Dey, N., Balas, V.E., Eds.; Academic Press: New York, NY, USA, 2020; Volume 6, pp. 1–35. [CrossRef]

15.　Kroll, M.; Kurinczuk, J.Q.M.; Hollowell, J. Ethnic disparity in risk of SIDS and other unexplained infant death is not due to deprivation; examining ethnic patterns may help to clarify aetiology. *J. Epidemiol. Community Health* **2018**, *72*, 911–918.

16.　Mitterauer, B. The gliocentric hypothesis of the pathophysiology of the sudden infant death syndrome (SIDS). *Med. Hypotheses* **2011**, *76*, 482–485. [CrossRef]

17.　Moon, R. SIDS and Other Sleep-Related Infant Deaths: Expansion of Recommendations for a Safe Infant Sleeping Environment. *Pediatrics* **2011**, *128*, 1030–1039. [CrossRef] [PubMed]

18.　O'Brien, M. SIDS—A developmental neurological perspective. Early Human Development. *Early Hum. Dev.* **1993**, *34*, 125–132. [CrossRef] [PubMed]

19.　Harpey, J.; Charpentier, C.; Paturneau-Jouas, M. Sudden Infant Death Syndrome and Inherited Disorders of Fatty Acid p-Oxidation. *Neonatology* **1990**, *58*, 70–80. [CrossRef] [PubMed]

20.　Álvarez Fumero, R. ¿Son los errores congénitos del metabolismo causa prevenible de muerte súbita? *Rev. Cuba. Pediatra* **2004**, *76*, 0–0.

21.　Chavez-Ocaña, S.; Bravata-Alcantara, J.; Sierra-Martinez, M. Errores innatos del metabolismo, una mirada a un tópico poco valorado. *Revista del Hospital Juárez de México* **2018**, *85*, 159–167.

22. Kumta, N. Inborn errors of metabolism (IEM)—An Indian perspective. *Indian J. Pediatr.* **2005**, *72*, 325–332. [CrossRef]

23. Gill, P.; van Zelm, M.; Muir, J.; Gibson, P. Review article: Short chain fatty acids as potential therapeutic agents in human gastrointestinal and inflammatory disorders. *Aliment. Pharmacol. Ther.* **2018**, *48*, 15–34. [CrossRef]

24. Jorda Lope, A. *Libro Blanco de la Muerte Subita Infantil*, 3th ed.; Grupo de Trabajo de Muerte Súbita Infantil-AEP:Madrid,España, 2013, pp.149–158, ISBN 978-84-15351-90-0.

25. Villoria, J.G.; Pajares, S.; López, R.M.; Marin, J.L.; Ribes, A. Neonatal screening for inherited metabolic diseases in 2016. In *Proceedings of the Seminars in Pediatric Neurology*; Elsevier: Amsterdam, The Netherlands, 2016; Volume 23, pp. 257–272.

26. Ogier de Baulny, H. Management and emergency treatments of neonates with a suspicion of inborn errors of metabolism. *Semin. Neonatol.* **2002**, *7*, 17–26. [CrossRef]

27. Ibdah, J. Fatty Acid Oxidation Defects as a Cause of SIDS, 2001. Available online: https://www.sleepreviewmag.com/sleep-health/prevailing-attitude/sleep-safety/fatty-acid-oxidation-defects-as-a-cause-of-sids/ (accessed on 26 March 2020).

28. Kononenko, I. Machine learning for medical diagnosis: History, state of the art and perspective. *Artif. Intell. Med.* **2001**, *23*, 89–109. [CrossRef]

29. Rodríguez-Ruiz, J.G.; Galván-Tejada, C.E.; Zanella-Calzada, L.A.; Celaya-Padilla, J.M.; Galván-Tejada, J.I.; Gamboa-Rosales, H.; Luna-García, H.; Magallanes-Quintanar, R.; Soto-Murillo, M.A. Comparison of Night, Day and 24 h Motor Activity Data for the Classification of Depressive Episodes. *Diagnostics* **2020**, *10*, 162. [CrossRef]

30. Alcalá-Rmz, V.; Zanella-Calzada, L.A.; Galván-Tejada, C.E.; García-Hernández, A.; Cruz, M.; Valladares-Salgado, A.; Galván-Tejada, J.I.; Gamboa-Rosales, H. Identification of diabetic patients through clinical and para-clinical features in Mexico: An approach using deep neural networks. *Int. J. Environ. Res. Public Health* **2019**, *16*, 381. [CrossRef] [PubMed]

31. Garcia-Ceja, E.; Osmani, V.; Mayora, O. Automatic stress detection in working environments from smartphones' accelerometer data: A first step. *IEEE J. Biomed. Health Inform.* **2015**, *20*, 1053–1060. [CrossRef]

32. Oriol, J.D.V.; Vallejo, E.E.; Estrada, K.; Peña, J.G.T.; The Alzheimer's Disease Neuroimaging Initiative Benchmarking machine learning models for late-onset alzheimer's disease prediction from genomic data. *BMC Bioinform.* **2019**, *20*, 1–17.

33. Kokol, P.; Završnik, J.; Blažun Vošner, H. Artificial intelligence and pediatrics: A synthetic mini review. *Pediatr. Dimens.* **2017**, *2*, 1–5. [CrossRef]

34. Mueller, M.; Almeida, J.S.R.; Wagner, C. Can Machine Learning Methods Predict Extubation Outcome in Premature Infants as well as Clinicians? *J. Neonatal Biol.* **2013**, *2*, 1–6. [CrossRef]

35. Wadhwani, S.I.; Hsu, E.K.; Shaffer, M.L.; Anand, R.; Ng, V.L.; Bucuvalas, J.C. Predicting ideal outcome after pediatric liver transplantation: An exploratory study using machine learning analyses to leverage Studies of Pediatric Liver Transplantation Data. *Pediatr. Transplant.* **2019**, *23*, e13554. [CrossRef]

36. NIH Common Fund's National Metabolomics Data Repository (NMDR) Website, t.M.W. SCFA Analysis in SIDS, Project ID ST000512. 2017. Available online: https://www.metabolomicsworkbench.org/data/DRCCMetadata.php?Mode=Project&ProjectID=PR000512 (accessed on 12 December 2019).

37. Galal, M.; Symonds, I.; Murray, H.; Petraglia, F.; Smith, R. Postterm pregnancy. *Facts Views Vis. ObGyn* **2012**, *4*, 175–187. [PubMed]

38. Joshi, A.V. *Machine Learning and Artificial Intelligence*, 1st ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2020.

39. Altman, D.G.; Bland, J.M. Diagnostic tests. 1: Sensitivity and specificity. *BMJ Br. Med. J.* **1994**, *308*, 1552. [CrossRef]

40. Hajian-Tilaki, K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp. J. Intern. Med.* **2013**, *4*, 627.

41. Nardo, L.G.; Gelbaya, T.A.; Wilkinson, H.; Roberts, S.A.; Yates, A.; Pemberton, P.; Laing, I. Circulating basal anti-Müllerian hormone levels as predictor of ovarian response in women undergoing ovarian stimulation for in vitro fertilization. *Fertil. Steril.* **2009**, *92*, 1586–1593. [CrossRef]

42. Yip, S.S.; Kim, J.; Coroller, T.P.; Parmar, C.; Velazquez, E.R.; Huynh, E.; Mak, R.H.; Aerts, H.J. Associations between somatic mutations and metabolic imaging phenotypes in non—Small cell lung cancer. *J. Nucl. Med.* **2017**, *58*, 569–576. [CrossRef] [PubMed]

43. Bianchi, A.R.; Ferreri, C.; Ruggiero, S.; Deplano, S.; Sunda, V.; Galloro, G.; Formisano, C.; Faraone Mennella, M.R. Automodification of PARP and fatty acid-based membrane lipidome as a promising integrated biomarker panel in molecular medicine. *Biomark. Med.* **2016**, *10*, 229–242. [CrossRef] [PubMed]

44. Caminha, T.C.; Ferreira, H.S.; Costa, N.S.; Nakano, R.P.; Carvalho, R.E.S.; Xavier, A.F., Jr.; Assunção, M.L. Waist-to-height ratio is the best anthropometric predictor of hypertension: A population-based study with women from a state of northeast of Brazil. *Medicine* **2017**, *96*, e5874. [CrossRef]

45. Mandrekar, J.N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* **2010**, *5*, 1315–1316. [CrossRef] [PubMed]

46. Schieda, N.; Dilauro, M.; Moosavi, B.; Hodgdon, T.; Gregory, O.C.; McInnes, M.D.F.; Flood, T.A. MRI evaluation of small (<4cm) solid renal masses: Multivariate modeling improves diagnostic accuracy for angiomyolipoma without visible fat compared to univariate analysis. *Eur. Radiol.* **2015**, *26*, 2242–2251. [CrossRef]

47. Wu, J.; Gong, G.; Cui, Y.; Li, R. Intratumor partitioning and texture analysis of dynamic contrast-enhanced (DCE)-MRI identifies relevant tumor subregions to predict pathological response of breast cancer to neoadjuvant chemotherapy. *J. Magn. Reson. Imaging* **2016**, *44*, 1107–1115. [CrossRef] [PubMed]

48. Asllani, I.; Habeck, C.; Scarmeas, N.; Borogovac, A.; Brown, T.R.; Stern, Y. Multivariate and univariate analysis of continuous arterial spin labeling perfusion MRI in Alzheimer's disease. *J. Cereb. Blood Flow Metab.* **2008**, *28*, 725–736. [CrossRef]

49. El-Hattab, A.W.; Sutton, V.R. Approach to inborn errors of metabolism in pediatrics. *Pediatr. Clin.* **2018**, *65*.10.1016/j.pcl.2018.01.001.

50. Optiz, E.; Gilbert-Barness, D.E.; Spicer, T.S.S. *Handbook of Pediatric Autopsy Pathology*, 2nd ed.; Springer: New York, NY, USA, 2013; pp. 653–673.

51. Byard, R.W.; Shipstone, R.A.; Young, J. Continuing major inconsistencies in the classification of unexpected infant deaths. *J. Forensic Leg. Med.* **2019**, *64*, 20–22. [CrossRef]

52. Loeber, J.G. Neonatal screening in Europe; the situation in 2004. *J. Inherit. Metab. Dis.* **2007**, *30*, 430–438. [CrossRef]

53. Wilcken, B.; Wiley, V.; Hammond, J.; Carpenter, K. Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N. Engl. J. Med.* **2003**, *348*, 2304–2312. [CrossRef]

54. Takiishi, T.; Fenero, C.I.M.; Câmara, N.O.S. Intestinal barrier and gut microbiota: Shaping our immune responses throughout life. *Tissue Barriers* **2017**, *5*, e1373208. [CrossRef]

55. Dai, X.; Yuan, T.; Zhang, X.; Zhou, Q.; Bi, H.; Yu, R.; Wei, W.; Wang, X. Short-chain fatty acid (SCFA) and medium-chain fatty acid (MCFA) concentrations in human milk consumed by infants born at different gestational ages and the variations in concentration during lactation stages. *Food Funct.* **2020**, *11*, 1869–1880. [CrossRef]