Muscle imaging in laminopathies: Synthesis study identifies meaningful muscles for follow-up

David Gómez-Andrés, Jordi Diaz-Manera, Aida Alejaldre, Irene Pulido-Valdeolivas, Laura González-Mera, Montse Olivé, Juan José Vilchez, Adolfo López de Munain, Carmen Paradas, Nuria Muelas, et al.

Muscle imaging in laminopathies: synthesis study identifies meaningful muscles for follow-up

Running title (26 characters): Muscle MRI progression in *LMNA*

Authors:

David Gómez-Andrés MD PhD[1], Jordi Díaz-Manera MD PhD[2], Aida Alejaldre MD PhD[2], Irene Pulido-Valdeolivas MD PhD[3], Laura González-Mera MD PhD[4,5], Montse Olivé MD PhD[5], Juan José Vilchez MD PhD[6], Adolfo López de Munain MD PhD[7], Carmen Paradas MD PhD[8], Nuria Muelas MD PhD[6], Ángel Sánchez-Montáñez MD[9], Alicia Alonso-Jimenez MD[2], Marta Gómez García de la Banda MD[10], Ivana Dabaj MD[10], Gisèle Bonne PhD[11], Francina Munell MD PhD[1], Robert Y Carlier MD[12] and Susana Quijano-Roy MD PhD[13].

1. Neuromuscular disorders Group. Child Neurology Unit. Hospital Universitari Vall d'Hebron. Vall d'Hebron Research Institute (VHIR). Barcelona, Spain.

2. Neuromuscular Disorders Unit. Neurology Department. Universitat Autònoma de Barcelona. CIBERER. Barcelona, Spain.

3. Center of Neuroimmunology. Institut d'Investigacions Biomèdiques August Pi i Sunyer. Hospital Clínic. Barcelona, Spain.

4. Department of Neurology. Hospital de Viladecans. Barcelona. Spain.

5. Institute of Neuropathology, Department of Pathology and Neuromuscular Unit, Department of Neurology, IDIBELL-Hospital de Bellvitge, Hospitalet de Llobregat, Barcelona, Spain

6. Department of Neurology, Hospital Universitari I Politècnic La Fe, CIBERER Valencia, Spain

7. Department of Neurology, Hospital Universitario Donostia, Donostia-San Sebastián, Spain. Neurosciences Area, Biodonostia Institute, CIBERNED. Donostia-San Sebastián, Spain

8. Neuromuscular Disorders Unit. Department of Neurology and Neurophysiology. Hospital Universitario Virgen del Rocío. Sevilla, Spain

9. Pediatric Radiology Department. Hospital Universitari Vall d'Hebron. Barcelona. Spain.

10. APHP, Neuromuscular Disorders Unit. Pediatric Department, CHU Paris IdF Ouest - Hôpital Raymond Poincaré, Garches, France

11. Sorbonne Université, Inserm UMRS974, Center for Research in Myology, Institut de Myologie, G.H. Pitié Salpêtrière, Paris, France.

12. APHP, Radiology Department. CHU Paris IdF Ouest - Hôpital Raymond Poincaré. Paris Saclay Universities, UVSQ University of Versailles, UMR 1179 INSERM Garches, France.

13. APHP, Neuromuscular Disorders Unit. Pediatric Department, CHU Paris IdF Ouest - Hôpital Raymond Poincaré, . Paris Saclay Universities, UVSQ University of Versailles, UMR 1179 INSERM. Garches, France

Corresponding author: David Gómez Andrés. Unitat de Neurologia Pediàtrica.

Hospital Universitari Vall d'Hebron. Passeig de la Vall d'Hebron 119-129. 08035

Barcelona. Spain. Email: david_gomez@vhebron.net

Disclosure of conflict of interest : None of the authors has any conflict of interest to disclose

Part of this work has been presented in 22th International Congress of the WMS

Abstract (149 words)

INTRODUCTION: Particular fibroadipose infiltration patterns have been recently described by muscle imaging in congenital and later-onset forms of *LMNA*-related muscular dystrophies (*LMNA*-RD).

METHODS: Scores for fibroadipose infiltration of 23 lower limb muscles in 34 patients with *LMNA*-RD were collected from heatmaps of two previous studies. Scoring systems were homogenized. Relationships between muscle infiltration and disease duration and age of onset were modeled with random forests.

RESULTS: The pattern of infiltration differs according to disease duration, but not to age of disease onset. The muscles whose progression best predicts disease duration were semitendinosus, biceps femoris long head, gluteus medius and semimembranosus.

DISCUSSION: In *LMNA*-RD, synthetic analysis of lower limb muscle infiltration did not find major differences between forms with different ages of onset, but allowed the identification of muscles with characteristic infiltration during disease progression. Monitoring of these specific muscles by quantitative MRI may provide useful imaging biomarkers in *LMNA*-RD.

Keywords: *LMNA*, imaging, magnetic resonance, biomarker, machine learning, laminopathy

Introduction

Muscle imaging has been increasingly recognized in the last decade as a valuable tool to facilitate the differential diagnosis in patients with hereditary muscular disorders. However, series are still small, and the relationships of imaging phenotypes with clinical data such as length of disease course, severity and age of onset require further study. It would be desirable to develop tools allowing analysis of integrated data from different series and enhancing the study of relationships between imaging and clinical data.

LMNA related muscular dystrophies (*LMNA*-RD) are a spectrum of muscular dystrophies due to mutations in the gene coding for Lamins A and C. Different forms have been described depending on age of onset and clinical presentation, and seem to constitute a continuum of phenotypes [1]. There is a congenital form (L-CMD) [2], an Emery-Dreifuss muscular dystrophy (EDMD) phenotype [3] and a form of limb girdle muscular dystrophy (LGMD1B)[4,5].

Several studies have been published with descriptions of signal abnormalities using T1-TSE sequences on magnetic resonance imaging (MRI) or computerized tomography (CT) scans [6-10]. More recently, two systematic analyses of muscle MRI have described the differences of fibroadipose infiltration in the muscle of *LMNA*-RD patients, one using whole body MRI in children with the congenital form [11] and the other using MRI and CT scans of lower limbs in adults with later onset forms – EDMD and LGMD1B [12]. Patterns of fatty infiltration have been systematically described. However, the course of fibroadipose substitution in different muscles during the course of the disease and the correspondence with the spectrum of different presentations had not been carefully studied.

Our objectives are 1) to demonstrate that the use of systematic muscle fatty infiltration scoring in imaging studies and their representation of results in heatmaps allows the collection of data from different series and the performance of synthesis studies (studies that pool data from various sources) based on individual data; 2) to detect, which muscles, if any, are differentially affected according to age at onset and which muscles, if any, are affected more as a result of disease duration in *LMNA*-related muscle disorders using previous published data. [11,12]

Methods

<u>Collection of data:</u>

From published heatmaps, [11,12] we collected fatty infiltration scores from 21 lower limb muscles that were imaged in both studies: gluteus maximus, gluteus medius, gluteus minimus, vastus lateralis, vastus medialis, vastus intermedius, rectus femoris, sartorius, gracilis, semimembranosus, semitendinosus, short head of biceps femoris, long head of biceps femoris, tibialis anterior, tibialis posterior, peroneus, extensor digitorum longus, flexor digitorum longus, soleus, medial head of gastrocnemius and lateral head of gastrocnemius. Gómez-Andrés et al.[11] scored the iliacus and psoas muscles independently whereas Díaz-Manera et al.[12] scored the iliopsoas in combination. To solve this, we calculated the median score of iliacus and psoas from Gómez-Andrés et al.[11] We also calculated the median score of the adductor brevis, longus and magnus to have an equivalent to the global scoring of adductor muscles performed by Díaz-Manera et al.[12] In general, the values of individual muscles within these two muscle groups in the work of Gómez-Andrés et al.[11] were very similar. We collected age of onset and calculated

the time of disease duration from the onset until imaging was performed in each subject. We used data from the 8 early-onset patients of Gómez-Andrés et al.[11] and 26 patients from Díaz-Manera et al.[12] (9 EDMD, 12 LGMD and 5 presymptomatic patients, in which disease duration was represented as a negative value. Ten patients were evaluated by means of lower limb CT imaging and 16 patients using lower limb MRI). Clinical features of the patients are shown in supplementary table 1. Research ethics committees approved both studies and informed consent was obtained from the research subjects or their guardians to perform MRI and analyzing the results.

Preparation of data: Muscle imaging scoring was performed using different scales in the two studies. Díaz-Manera et al.[12] used a modified Mercuri scale that quantifies muscle fatty infiltration from normal (0) to end-stage (4). Gómez-Andrés et al.[11] used a modification of Lamminen scale[13] that assesses fatty infiltration from normal (1) to complete fatty substitution (4). We proposed and applied a system of equivalencies between these two scales (see supplementary table 2). In order to estimate the impact of using this proposed system, one researcher (IPV) scored the pediatric cohort with the Mercuri scale used in the adult cohort. Agreement was high or complete in every muscle (see supplementary figure 1).

Machine learning approach: Random forests were used to model the relationship between the pattern of muscle infiltration and disease onset and disease duration. Random forests are computer algorithms that make predictions about one parameter (output) based on multiple parameters (inputs) after a training period[14,15]. In this case, we trained two random forests: one was developed to predict disease onset and the other to predict disease duration. Both random

forests used the same inputs: the scoring of muscle infiltration for the selected muscles. Random forests were considered the most appropriate tool because they are able to deal with situations with a large number of predicting variables and a low number of patients and they are good in processing ordinal data. As random forests are based on classification and regression trees, they make their predictions by means of splitting the database. So, it is possible to include data in which the numerical values are not exact but are present in one extreme of the distribution. This was important in our case because we needed represent negative but unknown values of disease duration in pre-symptomatic patients and negative but unknown values of disease onset in patients with prenatal onset. Disease duration in pre-symptomatic patients was represented by minus 5 years and prenatal onset, by minus 0.25 years. For disease onset, pre-symptomatic patients were excluded. Moreover, they provide measures of "importance" which are quantifications of the relevance of each input parameter in the prediction. We used package "randomforestSRC"[16] in R software to train random forests with variable selection using minimal depth methodology[17]. The goodness of fit of each model (how accurate the random forest was in its predictions) was assessed by Spearman's rho correlation coefficient between actual output parameter and the out-of-the-bag prediction made by random forest. These relationships were also represented by means of scatter plots. We used Breiman-Cutler permutation variable importance to detect muscles in which infiltration is related with the two outputs (disease duration and disease onset). We represented relationships between important infiltrations and the outputs by means of boxplots.

Results

The goodness of fit of the model for disease duration is acceptable (Spearman's rho = 0.612, figure 1A) while in the case of disease onset, it was very low (Spearman's rho = 0.005, figure 1B). This means that the pattern of muscle infiltration in lower limbs changes with the progression of the disease, but there are no significant differences in the global pattern of muscle infiltration according to the different ages at onset.

The most relevant muscles to predict disease duration for the random forests were semitendinosus, long head of biceps femoris, gluteus medius and semimembranosus (figure 1C). The rest of evaluated muscles were not selected as relevant by the variable selection algorithm. The relationships between disease duration and infiltration in each muscle could be either linear (meaning in this case that the higher the score of infiltration, the longer the disease duration along the whole spectrum of disease progression) or non-linear (meaning in this case that the relationship is only important for a part of the spectrum of disease progression). Semitendinosus infiltration (figure 2A) is related with disease duration in a non-linear manner. A non-infiltrated semitendinosus is markedly related with short time of disease duration (less than 10 years). Patients with disease longer than 20 years can have variable semitendinosus infiltration, but some degree of infiltration is likely to be present. Infiltration of long head of biceps (figure 2B) is linearly related with disease duration. Patients with less than 10 years of duration show no or mild infiltration in long head of biceps while patients with disease duration longer than 30 years show moderate to severe infiltration. Gluteus medius infiltration (figure 2C) is also linearly related: the more infiltrated gluteus medius is, the longer the disease duration tend to be in the patient.

Semimembranosus infiltration (figure 2D) shows a non-linear relationship with disease duration. Only the patients with longer duration (more than 25 years) show a consistently severe infiltration whereas patients with shorter times of disease duration tend to have less infiltration, but infiltration is not progressively related with disease duration for a duration of 10-20 years. Other muscles such as gluteus maximum and minimus, vastus intermedius and medialis, lateral gastrocnemius and tibialis posterior are also related with disease duration, but their importance is lower. Supplementary figure 2 show the relationships between infiltration scoring of these muscle and disease duration.

While global pattern of muscle infiltration is not related with age of onset, there are some muscles that could be detected as different according to the age of onset (figure 1D). For instance, adductor muscles tend to be severely infiltrated in early-onset patients while preserved in late-onset patients, and semitendinosus is only severely infiltrated in some early-onset patients (supplementary figure 3).

Discussion

Muscle imaging is gaining popularity as an approach to direct genetic studies for patients with suspicion of a hereditary muscle disorders because muscle changes rarely occur randomly but they often follow a specific pattern that is different from one mutated protein to other one[18]. The description of this pattern had usually been qualitative, based on subjective criteria that were difficult to quantify and replicate. The extended systematic use of semi-quantitative scores to describe the degree of muscle fatty infiltration in the different imaging techniques improved information gathering and sharing, which improved the description of specific muscle profiles for different hereditary muscle disorders[19]. In addition to

systematic scoring, multivariate descriptive techniques have been incorporated to improve the pattern identification and representation [11,12,18,20-22].

One of the best studied myopathies on muscle imaging is the group of muscle skeletal laminopathies or *LMNA*-RD. Several studies have described a consistent pattern[6,8,10], including two systematic analyses which were recently published on the congenital and adult forms. While the congenital series studied signal abnormalities using T1-TSE sequences on whole body MRI [11], the adult multicenter study in *LMNA*-related EDMD and LGMD1B collected data from muscle MRI and CT scans, all limited to the lower limbs[12].These studies found that *LMNA*-related muscle dystrophies show a characteristic pattern of muscle infiltration.

The present study shows, by the integration of data included in the two systematic studies using machine-learning techniques, that the pattern of involvement changes with the progression of the disease, but not with the age of onset. These results have several implications. They strengthen the clinical impression of a continuum of phenotypes in the same myopathy rather than different muscle diseases due to a same gene defect. On the other hand, the analysis identifies several muscles that are most clearly related to the progression of the myopathy (semitendinosus, long head of biceps, gluteus medius and semimembranosus muscles), meaning that they are strong candidates for future studies looking for biomarkers to be monitored by muscle quantitative MRI techniques. Our findings also support that including time of disease in muscle MRI interpretation could be important guide genetic studies. For example, if duration of disease is long, , a *LMNA*-related muscle disorder is unlikely if infiltration is lacking in the semitendinosus, semimembranosus, long head of biceps femoris or gluteus medius. In contrast, these muscles are expected to be spared or only mildly

infiltrated in patients with shorter disease duration. When machine learning is implemented for differential diagnosis in muscle MRI, disease duration may be also relevant to include as this may increase the diagnostic performance of the algorithm.

The lack of relationship with the type of *LMNA*-RD (form of onset) should be interpreted with caution. Clinical observations strongly suggest that data from cervical, periscapular and upper limb regions is very relevant in the distinction between the phenotypes with different age of onset (congenital *vs* later onset), but unfortunately this data is not yet available in EDMD and LGMD. In the lower limbs, the only muscles involved at a variable degree are adductor muscles and semitendinosus. These findings, especially those affecting the semitendinosus muscle, may also be due to the longer disease in some patients with earlier onset.

We believe that looking for targets for quantitative MRI by means of strategies of (re-)analyzing semiquantitative scoring is important because focusing on a limited number of muscles selected as candidates by these strategies could accelerate, simplify and reduce the cost of the time-consuming phase of image processing in quantitative MRI studies. Moreover, there may be some muscles that could be overlooked if a non-targeted approach is performed. For instance, the muscles that seem to be more related in our study have not been considered in previous quantitative studies for other muscle diseases[12,23,24]. However, this approach should be undertaken with caution for optimizing longitudinal quantitative MRI studies because the conclusions about the relationships of disease duration with muscle fatty infiltration are based on the assumption that cross-sectional data represent the longitudinal evolution of the disorder. Moreover, the pace of

progression in fatty infiltration may be not uniform between patients and a cross-sectional design does not permit evaluation of this.

We showed that relationships between the course of the disease and the progression of signal abnormalities can be non-linear. A non-linear relationship between disease duration and signal abnormality in some muscles implies that in these muscles, the pace of progression for fibroadipose infiltration is not homogeneous throughout the disease course. We speculate that the differing rates of fatty infiltration over time in some muscles may represent different pathogenic mechanisms or secondary changes due to disuse atrophy as a result of loss of mobility. This result is also particularly interesting if studies in *LMNA*-RD are designed using muscle MRI parameters as biomarkers. For instance, the semitendinosus seems to be resistant to infiltration in the early years of the disease, but shows a faster rate of infiltration 5-10 years into the disease course. Based on this, we could propose fatty infiltration of the semitendinosus as a good candidate biomarker for patients with a disease duration of 10-20 years, but a potentially poor biomarker for monitoring patients with *LMNA*-RD early in the disease course (no significant infiltration before 10 years of disease duration) and in very advanced cases (variably infiltrated). An alternative explanation for non-linear relationships could be that the ordinal scale to assess fibroadipose infiltration does not reflect accurate proportional increases (e.g. the increase in fatty infiltration between 1 and 2 may be not the same as the increase between 2 and 3) in some muscles.

Understanding the progression pattern in muscle degeneration in muscle dystrophies is also important for guiding future studies to better understand the underlying pathophysiologic mechanisms in each disease. At a given point in the

disease course, pathologic mechanisms may vary between muscles. For a given muscle, pathologic mechanisms may vary according to the point in the disease course. Knowing the pace of progression of the disease could allow to interpret findings in future studies or to improve the targeting of muscle biopsy in patients. Another contribution of this study is its methodological proposal. Due to the spread of systematic scoring and heatmap representation in muscle imaging, we expect increasing amounts of imaging data from individual patients. Synthesis and further analysis of these data are critical to hasten our understanding of imaging phenotypes in myopathies, and in producing hypotheses that could be confirmed with additional longitudinal and/or quantitative studies. In particular, these kinds of studies will permit the evaluation of the individual rates of progression and help to identify factors (e.g. functional status, age at onset, mutation type or others) that may influence it[25].

Although synthesis studies seems to be attractive approaches to combining data from different sources, they also have limitations [26]. In muscle MRI and in this particular study, the most important limitations derive from the use of several scoring techniques and the different selection of muscles. This highlights the need for guidelines and consensus in scoring. To overcome this limitation we incorporated random forests, a data mining techniques that compensates for this. We believe that the incorporation of new data mining techniques such as random forests could lead to improved differential diagnoses, and an improved understanding of the variability within diseases[27].

Figure 1. Upper row: Scatterplot of the prediction (y-axis) against the actual value (x-axis) of disease duration (A) and age at onset (B). **Lower row:** Importance values of infiltration of the selected muscles in random forests predicting disease duration (C) and age at onset (D).

Figure 2. Relationships between disease duration and scoring of infiltration in the most important muscles in the random forest prediction for duration. A. Semitendinosus. B. Long head of biceps femoris. C. Gluteus medius and D. Semimembranosus.

LIST OF ABBREVIATION
CT: Computerized tomography
EDMD: Emery-Dreifuss muscular dystrophy
L-CMD: LMNA-related congenital form
LGMD1B: limb girdle muscular dystrophy 1B
LMNA-RD: LMNA-related muscular dystrophies
MRI: Magnetic resonance imaging.

BIBLIOGRAPHY

1.   Maggi L, Carboni N, Bernasconi P. Skeletal Muscle Laminopathies: A Review of Clinical and Molecular Features. Cells 2016;5.
2.   Quijano-Roy S, Mbieleu B, Bonnemann CG, et al. De novo LMNA mutations cause a new form of congenital muscular dystrophy. Annals of neurology 2008;64:177-86.
3.   Bonne G, Quijano-Roy S. Emery-Dreifuss muscular dystrophy, laminopathies, and other nuclear envelopathies. Handbook of clinical neurology 2013;113:1367-76.
4.   Muchir A, Bonne G, van der Kooi AJ, et al. Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). Human molecular genetics 2000;9:1453-9.
5.   van der Kooi AJ, van Meegen M, Ledderhof TM, McNally EM, de Visser M, Bolhuis PA. Genetic localization of a newly recognized autosomal dominant limb-girdle muscular dystrophy with cardiac involvement (LGMD1B) to chromosome 1q11-21. American journal of human genetics 1997;60:891-5.
6.   Pasqualin LM, Reed UC, Costa TV, et al. Congenital muscular dystrophy with dropped head linked to the LMNA gene in a Brazilian cohort. Pediatric neurology 2014;50:400-6.
7.   Deconinck N, Dion E, Ben Yaou R, et al. Differentiating Emery-Dreifuss muscular dystrophy and collagen VI-related myopathies using a specific CT scanner pattern. Neuromuscular disorders : NMD 2010;20:517-23.
8.   Carboni N, Mura M, Marrosu G, et al. Muscle imaging analogies in a cohort of patients with different clinical phenotypes caused by LMNA gene mutations. Muscle & nerve 2010;41:458-63.
9.   Mercuri E, Clements E, Offiah A, et al. Muscle magnetic resonance imaging involvement in muscular dystrophies with rigidity of the spine. Annals of neurology 2010;67:201-8.
10.   Mercuri E, Counsell S, Allsop J, et al. Selective muscle involvement on magnetic resonance imaging in autosomal dominant Emery-Dreifuss muscular dystrophy. Neuropediatrics 2002;33:10-4.
11.   Gomez-Andres D, Dabaj I, Mompoint D, et al. Pediatric laminopathies: Whole-body magnetic resonance imaging fingerprint and comparison with Sepn1 myopathy. Muscle & nerve 2016;54:192-202.
12.   Diaz-Manera J, Alejaldre A, Gonzalez L, et al. Muscle imaging in muscle dystrophies produced by mutations in the EMD and LMNA genes. Neuromuscular disorders : NMD 2016;26:33-40.
13.   Lamminen AE. Magnetic resonance imaging of primary skeletal muscle diseases: patterns of distribution and severity of involvement. The British journal of radiology 1990;63:946-50.
14.   Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychological methods 2009;14:323-48.
15.   Breiman L. Random Forests. Machine Learning 2001;45:5-32.
16.   Ishwaran H, Kogalur UB. Random Forests for Survival, Regression, and Classification (RF-SRC), R package version 2.5.0. 2017.

17.   Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ, Lauer MS. High-Dimensional Variable Selection for Survival Data. Journal of the American Statistical Association 2010;105:205-17.

18.   Fischer D, Bonati U, Wattjes MP. Recent developments in muscle imaging of neuromuscular disorders. Current opinion in neurology 2016;29:614-20.

19.   Fischer D, Wattjes MP. Neuromuscular Imaging. New York: Springer; 2013.

20.   Hankiewicz K, Carlier RY, Lazaro L, et al. Whole-body muscle magnetic resonance imaging in SEPN1-related myopathy shows a homogeneous and recognizable pattern. Muscle & nerve 2015;52:728-35.

21.   Garibaldi M, Tasca G, Diaz-Manera J, et al. Muscle MRI in neutral lipid storage disease (NLSD). Journal of neurology 2017;264:1334-42.

22.   Figueroa-Bonaparte S, Segovia S, Llauger J, et al. Muscle MRI Findings in Childhood/Adult Onset Pompe Disease Correlate with Muscle Function. PloS one 2016;11:e0163493.

23.   Willcocks RJ, Rooney WD, Triplett WT, et al. Multicenter prospective longitudinal study of magnetic resonance biomarkers in a large duchenne muscular dystrophy cohort. Annals of neurology 2016;79:535-47.

24.   Willis TA, Hollingsworth KG, Coombs A, et al. Quantitative muscle MRI as an assessment tool for monitoring disease progression in LGMD2I: a multicentre longitudinal study. PloS one 2013;8:e70993.

25.   Rooney WD, Berlow Y, Forbes SC, et al. Statistical modeling of 5-year longitudinal data from a large cohort of Duchenne muscular dystrophy (DMD) subjects: an interim look at temporal characteristics of disease progression from the ImagingDMD study. Imaging in Neuromuscular Disease 2017: First International Conference on Imaging in Neuromuscular Disease; 2017; Berlin: Journal of Neuromuscular Diseases. p. S34-5.

26.   Cooper HM. Hypotheses and problems in research synthesis. In: Cooper HM, Hedges LV, Valentine JC, eds. The handbook of research synthesis and meta-analysis. 2nd ed. New York: Russell Sage Foundation; 2009:19-35.

27.   Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. The New England journal of medicine 2016;375:1216-9.