



Πανεπιστήμιο Πειραιώς – Τμήμα
Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Πληροφορική»

Τίτλος Διατριβής	“Διατροφικοί βιοδείκτες και μηχανική μάθηση για εφαρμογές εξατομικευμένης διατροφής και βελτιστοποίησης της ανθρώπινης υγείας” “Nutritional biomarkers and machine learning for personalized nutrition applications and health optimization”
Όνοματεπώνυμο Φοιτητή	Δημήτριος Παναγούλιας
Πατρώνυμο	Πέτρος
Αριθμός Μητρώου	ΜΠΠΛ/18056
Επιβλέπων	Γεώργιος Τσιχριντζής, Καθηγητής

Τριμελής Εξεταστική Επιτροπή

Γεώργιος Τσιχριντζής
Καθηγητής

Διονύσιος Σωτηρόπουλος
Επίκουρος Καθηγητής

Ιωάννης Τασούλας
Επίκουρος Καθηγητής

ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη - Abstract

1. Εισαγωγή	
1.α Σύνοψη	2
1.β Σύνοψη κεφαλαίων	3
1.γ Στόχοι	4
1.δ Σκοπιά	4
2. Γενικές Έννοιες	
2.1 Εξατομικευμένη Ιατρική	5
2.2 Βιοδείκτες	8
2.2 Τεχνολογία sequencing επόμενης γενιάς	14
2.3 Παχυσαρκία	17
2.4 διατροφικοί βιοδείκτες	17
2.4α Metabolomics -μεταβολικό προφίλ	17
2.4β Metabolites - μεταβολίτες	19
2.4γ Nutritional genomics	20
2.4δ Standard Biochemistry Profile - Βιοχημικό προφίλ	21
3. Νευρωνικά δίκτυα, αναγνώριση προτύπων και datasets	
3.1 Multi Layer Neural Network	25
3.2 Περιβάλλον υλοποίησης	26
3.2 Datasets και σχεδιασμός	28
4. Υλοποίηση - Αποτελέσματα	
4.1 Ανάλυση δεδομένων Α'	31
4.2 Νευρωνικό δίκτυο Α'	32
4.1 Ανάλυση δεδομένων Β'	37
4.2 Νευρωνικό δίκτυο Β'	39
5. Επίλογος	
5.1 Πρόληψη	43
5.2 Μοντελοποίηση	43
5.3 Αυτοματισμοί	43
5.4 Προοπτική	44
Πηγές και βιβλιογραφία	
Παράρτημα	47

Περίληψη - Abstract

Το δόγμα της "one size fits all" προσέγγισης έχει ξεπεραστεί στο πεδίο της διάγνωσης των ασθενειών και της διαχείρισης των ασθενών και ακολουθείται πλέον η λογική της εξατομικευμένης ιατρικής "personalized medicine". Οι βιοδείκτες αποτελούν τις βασικές μεταβλητές στην έρευνα και την ανάπτυξη νέων μεθόδων εκπαίδευσης μοντέλων πρόγνωσης στο επιστημονικό πεδίο της μηχανικής μάθησης και της δημιουργίας ιατρικών εφαρμογών τεχνητής νοημοσύνης.

Σημαντικοί βιοδείκτες οι οποίοι σχετίζονται με τον μεταβολισμό είναι οι μεταβολίτες. Metabolomics είναι η συστηματική μελέτη των μοναδικών χημικών δακτυλικών αποτυπωμάτων (metabolites) που αφήνουν πίσω συγκεκριμένες κυτταρικές διεργασίες. Το μεταβολικό προφίλ μπορεί να δώσει ένα στιγμότυπο της φυσιολογίας του κυττάρου, και κατ' επέκταση τα metabolomics παρέχουν μια άμεση «λειτουργική ανάγνωση της φυσιολογικής κατάστασης» ενός οργανισμού.

Ο στόχος της παρούσας εργασίας είναι να αποτυπωθεί το γενικό διάγραμμα αξιολόγησης διατροφικών βιοδεικτών, να διερευνηθεί η δυνατότητα πρόβλεψης του δείκτη μάζας και να αναγνωριστούν διαιτητικά πρότυπα με τη χρήση νευρωνικών δικτύων.

The doctrine of the "one size fits all" approach has been overcome in the field of disease diagnosis and patient management and has been replaced by a more per patient approach known as personalized medicine. Bio markers are the key variables in the research and development of new methods of training prognostic models and neural networks in the scientific field of machine learning and artificial intelligence.

Important biomarkers related to metabolism are the metabolites. Metabolomics is the systematic study of unique chemical fingerprints that are left behind by specific cellular processes. The metabolic profile can provide a snapshot of cell physiology, and by extension metabolomics provide a direct "functional reading of the physiological state" of an organism.

The goal of this paper is to formulate a general evaluation chart of the nutritional biomarkers, to investigate how to best predict body mass index and to discover dietary patterns with the deployment of neural networks.

1. Εισαγωγή

1.α Σύνοψη

Το δόγμα της “one size fits all” προσέγγισης έχει ξεπεραστεί στο πεδίο της διάγνωσης των ασθενειών και της διαχείρισης των ασθενών και ακολουθείται πλέον η λογική της εξατομικευμένης ιατρικής “personalized medicine”. Οι βιοδείκτες αποτελούν τις βασικές μεταβλητές στην έρευνα και την ανάπτυξη νέων μεθόδων εκπαίδευσης μοντέλων πρόγνωσης στο επιστημονικό πεδίο της μηχανικής μάθησης και της δημιουργίας ιατρικών εφαρμογών τεχνητής νοημοσύνης.

Οι βιοδείκτες όμως **δεν αποτελούν μια καινούρια πραγματικότητα** καθώς υπάρχουν και χρησιμοποιούνται στην ιατρική επιστήμη τα τελευταία 20 χρόνια και μαζί τους έχει αναδυθεί η έννοια της εξατομικευμένης ιατρικής (personalised medicine). Γιατί είναι όμως σημαντικοί και πως μπορούν να μας βοηθήσουν στην καλύτερη διαχείριση των ασθενών και την πιο αποτελεσματική αντιμετώπιση του καρκίνου? Η διερεύνηση και ταυτοποίηση νέων βιοδεικτών θεωρητικά, θα μας βοηθά στο μέλλον να προβλέπουμε και να αντιμετωπίζουμε πιο γρήγορα και πιο αποτελεσματικά τις διάφορες ασθένειες και καταστάσεις της ανθρώπινης υγείας, και αυτό το γεγονός τους καθιστά τόσο σημαντικούς.

Παράλληλα υπάρχουν **σημαντικά ζητήματα** κοινωνικά, επιστημονικά, ηθικά και οικονομικά τα οποία περιβάλλουν αυτό το πεδίο καθώς πρώτον **η πολυπλοκότητα της βιολογίας του ασθενειών και του ανθρώπινου σώματος** μπορεί να περιορίσει την κλινική αποτελεσματικότητα της χρήσης των βιοδεικτών. Δεύτερον **το υψηλό κόστος θεραπειών όπως του καρκίνου σε συνδυασμό με τους περιορισμένους πόρους** του οικονομικού συστήματος δημιουργούν αντικειμενικά προβλήματα στο πως οι πόροι θα διοχετευτούν σε ποιες ομάδες, με τι κριτήρια, με τι επιπτώσεις και με τι αποτελέσματα.

Καθώς η έρευνα είναι δυναμική και τα σχετικά δεδομένα συνεχώς αλλάζουν τα παραπάνω ζητήματα **θα πρέπει διαρκώς να εξετάζονται και να επαναξιολογούνται**.

Σημαντικοί βιοδείκτες οι οποίοι σχετίζονται με τον μεταβολισμό είναι τα metabolites. Metabolomics είναι η συστηματική μελέτη των μοναδικών χημικών δακτυλικών αποτυπωμάτων (metabolites) που αφήνουν πίσω συγκεκριμένες κυτταρικές διεργασίες. Το μεταβολικό προφίλ μπορεί να δώσει ένα στιγμότυπο της φυσιολογίας του κυττάρου, και κατεπέκταση τα metabolomics παρέχουν μια άμεση «λειτουργική ανάγνωση της φυσιολογικής κατάστασης» ενός οργανισμού.

Η παχυσαρκία αντιπροσωπεύει μείζον πρόβλημα υγείας, όχι μόνο στη Δύση, αλλά όλο και περισσότερο στις χώρες με χαμηλό και μεσαίο εισόδημα. Προκειμένου να αναπτυχθούν επιτυχημένες στρατηγικές για την απώλεια βάρους, είναι σημαντικό να κατανοηθεί η μοριακή παθογένεση της αλλαγής βάρους. Ορισμένα μονοπάτια, όπως το οξειδωτικό στρες αλλά και το θεμελιώδες ρυθμιστικό πλαίσιο της ινσουλίνης, έχουν εμπλακεί στην αύξηση βάρους και στη ρύθμιση των ενεργειακών δαπανών. Επιπλέον, έχει επισημανθεί ο ρόλος των μεταβολιτών που παράγονται από το εντερικό μικρόβιο, ιδιαίτερα των λιπαρών οξέων βραχείας αλυσίδας.

1.β Σύνοψη κεφαλαίων

Παρακάτω θα παρατεθεί η σύνοψη κάθε κεφαλαίου ξεχωριστά και θα αναφερθούν επιγραμματικά οι βασικές έννοιες προς διευκόλυνση του αναγνώστη.

Στο **κεφάλαιο 2** θα αναλυθούν οι βασικές έννοιες οι οποίες θα χρησιμοποιηθούν και αφορούν την καλύτερη κατανόηση του εξεταζόμενου αντικειμένου.

- **Εξατομικευμένη ιατρική:** Η λογική της εξατομίκευσης έχει ως στόχο να δώσει καίριες λύσεις σε σχέση με την προγενέστερη προσέγγιση, μια λύση δηλαδή για όλους, που συχνά οδηγούσε σε ανεπαρκή ή υπερβολική θεραπεία των ασθενών και σε συνακόλουθο κίνδυνο ανεπιθύμητων ενεργειών.
- **Βιοδείκτες:** Ένας τρόπος για την εξατομίκευση των θεραπειών μπορεί να γίνει μέσω των βιοδεικτών. Οι βιοδείκτες είναι μόρια (πρωτεΐνες ή αντισώματα) ή βιοχημικές μεταβολές (γονιδιακές εκφράσεις και μεταλλάξεις) που βρίσκονται σε ιστούς, στο αίμα ή σε άλλα σωματικά υγρά ασθενών και υποδηλώνουν την παρουσία κάποιας συγκεκριμένης ασθένειας, πάθησης ή μετάλλαξης στο σώμα.
- **Παχυσαρκία:** Αποτελεί σημαντικό πρόβλημα δημόσιας υγείας. Η υψηλότερη πρόσληψη ενέργειας από τη συνολική ενέργεια που δαπανάται για μεγάλο χρονικό διάστημα οδηγεί σε παχυσαρκία. Αν και πολλοί άνθρωποι ειδικά στις δυτικές χώρες είναι υπέρβαροι ή παχύσαρκοι, ένα σημαντικό ποσοστό ατόμων με κανονικό βάρος δεν γίνονται ποτέ υπέρβαροι ή παχύσαρκα, αντανακλώντας εν μέρει τη μεγάλη διατομική διακύμανση της περίσσειας θερμιδικής πρόσληψης.
- **Metabolomics - διατροφικοί βιοδείκτες:** Η συστηματική μελέτη των μοναδικών χημικών δακτυλικών αποτυπωμάτων που αφήνουν πίσω συγκεκριμένες κυτταρικές διεργασίες. Το μεταβολικό προφίλ μπορεί να δώσει ένα στιγμότυπο της φυσιολογίας Στου κυττάρου, και κατεπέκταση τα metabolomics παρέχουν μια άμεση «λειτουργική ανάγνωση της φυσιολογικής κατάστασης» ενός οργανισμού
- **Metabolites - διατροφικοί βιοδείκτες:** Οι μεταβολίτες είναι τα υποστρώματα, τα ενδιάμεσα και τα προϊόντα του μεταβολισμού. Στο πλαίσιο της μεταβολικής, ένας μεταβολίτης ορίζεται συνήθως ως οποιοδήποτε μόριο σε μέγεθος μικρότερο από 1,5 kDa. Ωστόσο, υπάρχουν εξαιρέσεις σε αυτό ανάλογα με το δείγμα και τη μέθοδο ανίχνευσης. Για παράδειγμα, μακρομόρια όπως λιποπρωτεΐνες και λευκωματίνη ανιχνεύονται αξιόπιστα σε μελέτες μεταβολικής με βάση το NMR στο πλάσμα του αίματος.
- **Nutritional genomics - διατροφικοί βιοδείκτες:** Οι στόχοι της διατροφικής γονιδιωματικής είναι οι εξής: Προσδιορισμός γενετικών παραλλαγών που μπορεί να είναι σημαντικές στην κατανόηση των γενετικών αποκρίσεων στη διατροφή.
- **Standard Biochemistry Profile:** Οι βιοχημικές δοκιμές χρησιμοποιούνται για τον εντοπισμό βακτηριακών ειδών με τη διαφοροποίησή τους βάσει βιοχημικών δραστηριοτήτων. Η διαφορά στον μεταβολισμό των πρωτεϊνών και των λιπών, του μεταβολισμού των υδατανθράκων, της παραγωγής ενζύμων, της ικανότητας χρήσης ενώσεων κ.λπ.
- **Νευρωνικά δίκτυα:** Τα δίκτυα πολλαπλών επιπέδων επιλύουν το πρόβλημα ταξινόμησης για μη γραμμικά σύνολα χρησιμοποιώντας κρυφά στρώματα, των οποίων οι νευρώνες δεν συνδέονται άμεσα με την έξοδο. Τα επιπρόσθετα κρυμμένα

στρώματα μπορούν να ερμηνευθούν γεωμετρικά ως επιπρόσθετα υπερ-επίπεδα, τα οποία ενισχύουν την χωρητικότητα χωρισμού του δικτύου.

Στο **κεφάλαιο 3** θα αναλυθούν τα τεχνικά εργαλεία και τα datasets τα οποία θα χρησιμοποιηθούν για την δημιουργία των νευρωνικών δικτύων, οι σχετικές μέθοδοι, οι βιβλιοθήκες και τα προγραμματιστικά frameworks.

Στο **κεφάλαιο 4** θα αποτυπωθούν τα αποτελέσματα του εκτελέσιμου προγράμματος καθώς και ο μέρος του πηγαίου κώδικα.

Το **κεφάλαιο 5** περιέχει κάποιες προοπτικές σε σχέση με τα πορίσματα της συγκεκριμένης εργασίας.

1.γ Στόχοι

Ο στόχος της παρούσας εργασίας είναι να αποτυπωθεί το γενικό διάγραμμα αξιολόγησης διατροφικών βιοδεικτών, να διερευνηθεί η δυνατότητα πρόβλεψης του δείκτη μάζας και να αναγνωριστούν διαιτητικά πρότυπα με τη χρήση νευρωνικών δικτύων.

Κατόπιν θα επιλεγεί το μοντέλο-νευρωνικό δίκτυο και τα χαρακτηριστικά τα οποία δίνουν το καλύτερο αποτέλεσμα πρόβλεψης με το μικρότερο δυνατό κόστος.

1.δ Σκοπιά

Η ανάλυση θα γίνει μέσω της θεωρητικής περιγραφής του προβλήματος, την μοντελοποίηση του μέσα από αυτοματοποιημένη εφαρμογή αλγορίθμων μηχανικής μάθησης και της αναγνώρισης προτύπων. Για να γίνει αυτό έχει χρησιμοποιηθεί εκτενής βιβλιογραφία για την κατανόηση του προς εξέταση αντικείμενου. Το πρόβλημα έχει επιλυθεί σε περιβάλλον python με τη χρήση κατάλληλων βιβλιοθηκών.

2. Γενικές Έννοιες

2.1 Εξατομικευμένη Ιατρική

Η αποτελεσματικότητα του ιατρικού συστήματος εξαρτάται από τη δυνατότητα του να προβλέπει και να ανταποκρίνεται στις ανάγκες των πολιτών για ιατροφαρμακευτική περίθαλψη εξασφαλίζοντας χαμηλό κόστος και ποιότητα αξιοποιώντας τους διαθέσιμους πόρους - χρηματοοικονομικούς, τεχνολογικούς και ανθρώπινους (Bosanquet & Sikora 2004). Η ιατρική επιστήμη όπως και η οικονομική βασίζεται σε εμπειρικά και πραγματικά δεδομένα για να λάβει αποφάσεις και για να διαμορφώσει μοντέλα ώστε να είναι δυνατή η παροχή κοινωνικά δίκαιων και ωφέλιμων λύσεων.

Τα τελευταία 20 χρόνια ο κλάδος της ιατρικής έχει υποστεί μια σημαντική αλλαγή. Έχει ξεπεραστεί το δόγμα της μίας θεραπείας για όλους “one size fits all” προσέγγισης και ακολουθείται η λογική της εξατομικευμένης ιατρικής “personalized medicine”. Η λογική της εξατομικεύσης έχει ως στόχο να δώσει καίριες λύσεις στην προηγούμενη προσέγγιση που συχνά οδηγούσε σε ανεπαρκή ή υπερβολική θεραπεία των ασθενών και σε συνακόλουθο κίνδυνο ανεπιθύμητων ενεργειών. Στόχος της ιατρικής σήμερα είναι να προσαρμόζονται οι θεραπείες σε υποομάδες ασθενών που μοιράζονται παρόμοια γενετικά χαρακτηριστικά.(EC 10/3/2016)

Οι προσεγγίσεις συγκριτικής αποτελεσματικότητας για τη δημιουργία αποδείξεων, συμπεριλαμβανομένων των μετρήσεων των αποτελεσμάτων της υγείας, της ποιότητας ζωής, των οικονομικών αναλύσεων, της μοντελοποίησης αποφάσεων και των πραγματικών κλινικών δοκιμών, μπορούν να χρησιμοποιηθούν για να παρέχουν στους ενδιαφερόμενους μια σειρά πληροφοριών για την ενημέρωση σχετικά με τη θεραπεία, τις κατευθυντήριες γραμμές, την κάλυψη και τις αποφάσεις αποζημίωσης.

Η Ευρωπαϊκή Επιτροπή υπό την λουξεμβουργιανή Προεδρία έκανε την εξατομικευμένη ιατρική μια από τις προτεραιότητές της στον τομέα της υγείας και ενέκρινε στις 7 Δεκεμβρίου 2015 (EC. 2015). τα συμπεράσματα του Συμβουλίου στον τομέα αυτό. Σύμφωνα με αυτά “η εξατομικευμένη ιατρική αναφέρεται σε ένα ιατρικό μοντέλο που χρησιμοποιεί τον χαρακτηρισμό των φαινοτύπων και των γονότυπων των ατόμων (π.χ. μοριακό προφίλ, ιατρική απεικόνιση, δεδομένα για τον τρόπο ζωής) για την προσαρμογή της σωστής θεραπευτικής στρατηγικής για τον κατάλληλο άνθρωπο την κατάλληλη στιγμή ή / να καθορίσει την προδιάθεση για ασθένεια και / ή να παράσχει έγκαιρη και στοχοθετημένη πρόληψη. Στο πλαίσιο αυτό, το Συμβούλιο κάλεσε τα κράτη μέλη και την Επιτροπή μέσω της STAMP να αναλύσουν τα θέματα που σχετίζονται με την εφαρμογή της νομοθεσίας της Ευρωπαϊκής Ένωσης στον τομέα των φαρμάκων με στόχο τον εντοπισμό τρόπων μεγιστοποίησης της αποτελεσματικής χρήσης των υφιστάμενων ρυθμιστικών μέσων της Ευρωπαϊκής Ένωσης και περαιτέρω βελτίωση της ασφαλούς και έγκαιρης πρόσβαση στα φάρμακα για τους ασθενείς, συμπεριλαμβανομένων των καινοτόμων φαρμάκων · και να συνεχίσει να παρακολουθεί την πρόοδο του πιλοτικού προγράμματος προσαρμοστικής οδού που έχει αναλάβει ο Ευρωπαϊκός Οργανισμός Φαρμάκων και το δυναμικό του να επιτρέψει την έγκαιρη έγκριση ενός φαρμάκου για χρήση σε έναν καλά καθορισμένο πληθυσμό ασθενών με υψηλή ιατρική ανάγκη.”(STAMP, march 2016).

Στην εξέλιξη αυτής της πρωτοβουλίας θεμελιώθηκε ο οργανισμός ERA perMed ο οποίος αποτελείται από 32 ερευνητικούς και επενδυτικούς οργανισμούς από 23 χώρες και τη συγχρηματοδότηση της Ευρωπαϊκής επιτροπής. Σκοπός του αποτελεί η ευθυγράμμιση και χρηματοδότηση των ερευνητικών στρατηγικών, η ενίσχυση της

ανταγωνιστικότητας των ευρωπαϊκών συμμετεχόντων στην προσπάθεια της εξατομικευμένης ιατρικής και την ενδυνάμωση των αποτελεσμάτων της.(<http://www.eraermed.eu>).

Για την καλύτερη εκτέλεση και αναγνώριση των σχετικών αναγκών αναπτύχθηκε ένα θεσμικό πλαίσιο το οποίο αποτελείται από 2 πυλώνες.

α) Της κλινικής έρευνας.

Η εξατομίκευση των θεραπειών υπαγορεύει την ενοποίηση όλων των δεδομένων υγείας (ιατρικά, απεικονιστικά, τρόπος ζωής) ώστε να μπορούν να πραγματοποιηθούν ουσιώδης και αποτελεσματικές παρεμβάσεις.

Στο κομμάτι των κλινικών δοκιμών υπάρχουν αρκετές προκλήσεις οι οποίες θα πρέπει να αξιολογηθούν για χάρη της καινοτομίας. Πρέπει να εξεταστεί μια σειρά νέων, πιο ευέλικτων εναλλακτικών κλινικών δοκιμών και προσεγγίσεων παραγωγής δεδομένων, στατιστικών μεθόδων και εργαλείων ανάλυσης.

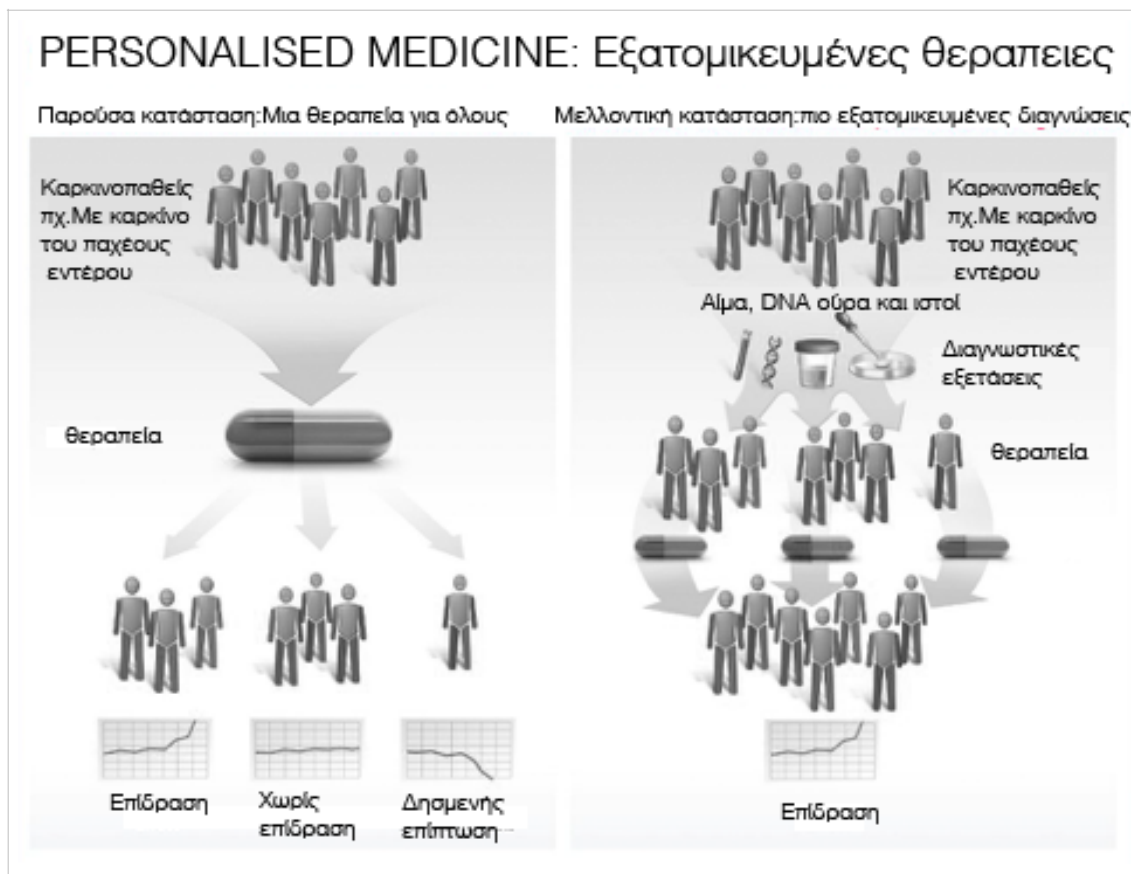
β) Της αδειοδότησης.

Η κοινή ανάπτυξη φαρμάκων και ιατροτεχνολογικών προϊόντων φαίνεται πολύπλοκη στην Ευρωπαϊκή Ένωση λόγω των διαφορετικών ρυθμιστικών συστημάτων.

Αξίζει να σημειωθεί ότι αυτά τα φάρμακα είναι δυνητικά επιλέξιμα για έναν αριθμό ρυθμιστικών οδών / κινήτρων, ήτοι η χρήση της άδειας κυκλοφορίας υπό όρους, η έγκριση σε εξαιρετικές περιπτώσεις και σε περιπτώσεις όπου η τυποποιημένη ανάπτυξη δεν είναι εφικτή, η ταχεία αξιολόγηση ή η δεκαετής αποκλειστικότητα μέσω της ορφανικής ονομασίας. Η διεθνής εναρμόνιση των κανονιστικών απαιτήσεων θα μπορούσε επίσης να διευκολύνει την αποδοχή προϊόντων πέραν των συνόρων και μπορεί να προωθηθεί.

Εξίσου σημαντική είναι η εκπαίδευση και η ενημέρωση των επαγγελματιών υγείας στο πεδίο της εξατομικευμένης ιατρικής. Σε αυτό το πλαίσιο θεωρείται ιδιαίτερα σημαντική η περιγραφή των χαρακτηριστικών των σχετικών φαρμάκων, η προτροπή για γενετικό έλεγχο και η ακριβής διατύπωση των ζητούμενων μεταλλάξεων για την αποτελεσματική χρήση των φαρμάκων. Για παράδειγμα, ο μοριακός χαρακτηρισμός των παθογόνων παραγόντων μπορεί να είναι σημαντικός πριν συνταγογραφηθεί το σωστό αντιβιοτικό στους ασθενείς και αποφευχθεί η επακόλουθη αντίσταση σε πολλαπλά φάρμακα.

Σε όλη αυτή τη διαδικασία είναι απαραίτητος ο συγχρονισμός του φαρμακευτικού, του ιατρικού κλάδου και των εποπτικών φορέων, ο έλεγχος της αποτελεσματικότητας των φαρμάκων και οι γρήγορες παρεμβάσεις για αλλαγές και βελτιώσεις όταν αυτό κρίνεται απαραίτητο.



Εικόνα 2.α Εξατομικευμένες θεραπείες παρόν και μέλλον

Πηγή: Jakka & Rossbach 2013

Διαφορετικοί άνθρωποι ανταποκρίνονται διαφορετικά στην ίδια θεραπεία, ενώ μία θεραπεία επιφέρει το επιθυμητό αποτέλεσμα σε μία ομάδα ασθενών δεν αλλάζει καθόλου την κατάσταση άλλων ομάδων ή οδηγεί σε δυσμενείς συνέπειες όπως βλέπουμε και στον **εικόνα 2α**. Ο λόγος είναι ότι η γενετική σύνθεση και το μεταβολικό προφίλ κάθε ασθενή επηρεάζει την επίδραση ενός φαρμάκου. Η εξατομικευμένη ιατρική λαμβάνει υπόψη αυτά τα μεμονωμένα πρότυπα κυτταρικών και μεταβολικών παραγόντων στη φάση της διάγνωσης. Η διάγνωση των βιολογικών δεικτών διαχωρίζει τους ασθενείς σε ομάδες με παρόμοια χαρακτηριστικά και παρέχει πληροφορίες για την καλύτερη ατομική θεραπεία.

Μέρος των ευρύτερων θετικών επιδράσεων που θα προκύψουν μέσω της εξατομικευμένης ιατρικής αποτελούν η δυνατότητα πιο αποτελεσματικών παρεμβάσεων σε τοπικούς πληθυσμούς βάση των ιδιαίτερων κοινωνικών τους χαρακτηριστικών. Η καλύτερη κατανόηση της επίπτωσης της ηλικίας, του φύλου, της εθνικότητας και του τρόπου ζωής θα επιτρέψει την πιο έξυπνη και αποτελεσματική διαχείριση των περιορισμένων διαθέσιμων πόρων. Αυτό προϋποθέτει τη δημιουργία συνεργιών μεταξύ διαφορετικών επιστημονικών κλάδων και την ενεργή συμμετοχή των ασθενών. Η γοργή ανάπτυξη της τεχνολογίας παρέχει και θα συνεχίσει να παρέχει σε μεγαλύτερο βαθμό το κατάλληλο περιβάλλον για την βελτίωση των παροχών υγείας. Οι γονιδιωματικές τεχνολογίες ήδη αποτελούν σημαντικό κομμάτι της σύγχρονης ιατρικής.

Η αποφυγή δοκιμών και σφαλμάτων μέσω της καλύτερης και πραγματικά εξατομικευμένης ιατρικής, η χρήση γενετικών προφίλ για τον εντοπισμό του καλύτερου δυνατού φαρμάκου και θεραπείας για ένα συγκεκριμένο ασθενή και η μείωση των ανεπιθύμητων ενεργειών συγκαταλέγονται στους βασικούς εξατομικευμένους ιατρικούς στόχους που θα ωφελήσουν όχι μόνο τους ασθενείς. Η προληπτική, εξατομικευμένη και

συμμετοχική ιατρική θα βοηθήσει στον εντοπισμό του κατάλληλου φαρμάκου για τον κατάλληλο ασθενή την κατάλληλη στιγμή, αποφεύγοντας τη συνταγογράφηση δαπανηρών και αναποτελεσματικών φαρμάκων και αποτρέποντας δυνητικά επιβλαβείς παρενέργειες(εικόνα 2.α). Σε αυτό το σημείο είναι σημαντικό να γίνουν οι εξής παραδοχές.

- A. Η ιατρική είναι μια βασισμένη στην επιστήμη της πληροφορίας.
- B. Μια συστηματική προσέγγιση είναι άρρηκτα συνδεδεμένη με τη μελέτη της τεράστιας πολυπλοκότητας των ασθενειών και της ανάλυσης των ασθενειών,
- C. Τα μαθηματικά και η συνεχώς αυξημένη υπολογιστική ισχύς θα επιτρέψουν την ανάλυση χιλιάδων δεδομένων με τρομερά υψηλές ταχύτητες.

Η ευρεία εφαρμογή εξατομικευμένων ιατρικών προσεγγίσεων έχει δύο σημαντικές προκλήσεις. Η μία είναι η σύνθετη βιολογία του ανθρώπινου σώματος και τα περίπλοκα (μακρό - μικρό) οικονομικά που αφορούν τον κλάδο της ιατρικής. Το δικαίωμα στην υγεία είναι παγκόσμιο και θεμελιώδες. Οι περιορισμένοι πόροι όμως, το υψηλό κόστος των φαρμάκων και των ιατρικών εφαρμογών, το μεγάλο οικονομικό και χρονικό κόστος των ερευνητικών δράσεων αποτελούν συχνά εμπόδιο για την γενικευμένη εφαρμογή τέτοιων σύγχρονων πρακτικών και τη εγκαθίδρυση ενός νέου ιατρικού δόγματος. Είναι σημαντική για αυτό το λόγο η κατανόηση των θετικών επιδράσεων της εξατομικευμένης ιατρικής από τους ασθενείς και η συμμετοχή αυτών στις αποφάσεις. Οι μεγαλύτερες προκλήσεις είναι οικονομικού χαρακτήρα και όχι ιατρικού.

2.2 Βιοδείκτες

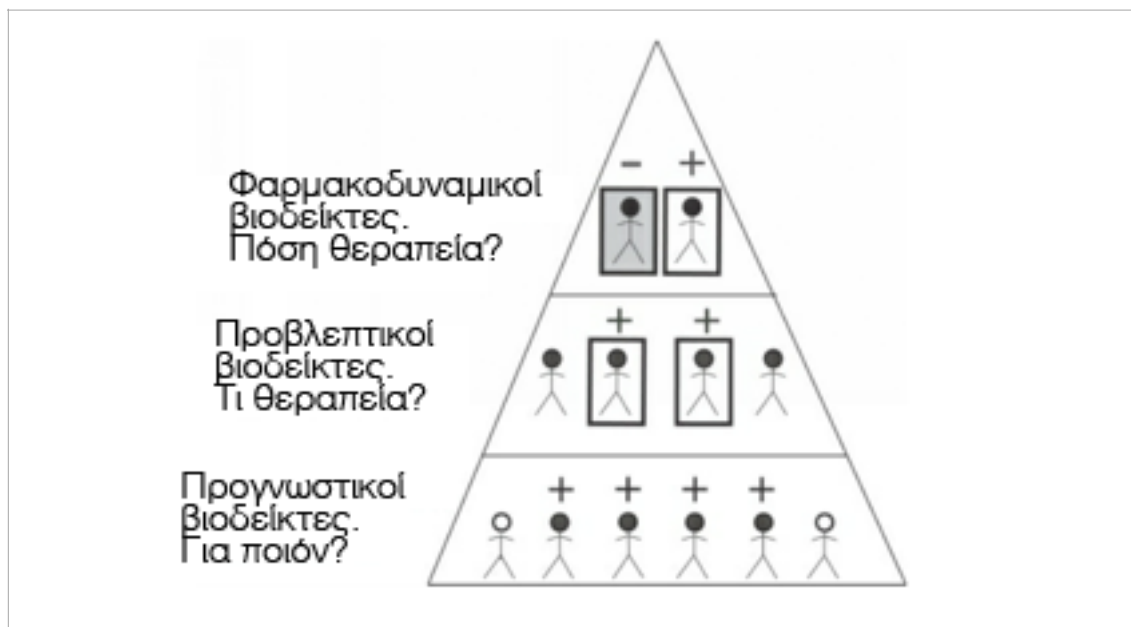
Ένας τρόπος για την εξατομίκευση των θεραπειών μπορεί να γίνει μέσω των βιοδεικτών. Οι βιοδείκτες είναι μόρια (πρωτεΐνες ή αντισώματα) ή βιοχημικές μεταβολές (γονιδιακές εκφράσεις και μεταλλάξεις) που βρίσκονται σε ιστούς, αίμα ή άλλα σωματικά υγρά ασθενών και υποδηλώνουν την παρουσία κάποιας συγκεκριμένης ασθένειας, πάθησης ή μετάλλαξης στο σώμα.

Οι βιοδείκτες μπορεί να είναι προγνωστικοί ή/και προβλεπτικοί. Χρησιμοποιούνται για να παρακολουθούν και να προβλέπουν την κατάσταση ή την εξέλιξη της υγείας των ασθενών. Χρησιμοποιούνται για την αξιολόγηση της ακολουθούμενης θεραπείας και την ανταπόκριση του ασθενή σε συγκεκριμένη φαρμακευτική αγωγή.(Romeo & Giamberini, 2012)

Η διερεύνηση και ταυτοποίηση νέων βιοδεικτών θεωρητικά, θα συνδράμει να προβλέπουμε και να αντιμετωπίζουμε πιο γρήγορα και πιο αποτελεσματικά τις διάφορες εκδηλώσεις του καρκίνου, και αυτό το γεγονός τους καθιστά τόσο σημαντικούς. Αυτή η υπόσχεση έχει οδηγήσει την Ευρωπαϊκή Ένωση και την Αμερικάνικη κυβέρνηση να δαπανά σημαντικά κονδύλια στην διαρκή έρευνα καρκινικών βιοδεικτών καθώς **μόνο ένας μικρός αριθμός σήμερα έχει κλινική εφαρμογή.**

Το 1998, η ομάδα εργασίας του Αμερικανικού Εθνικού Ινστιτούτου Υγείας καθόρισε ως βιοδείκτη «το χαρακτηριστικό που μετράται αντικειμενικά και αξιολογείται ως δείκτης φυσιολογικών βιολογικών διεργασιών, παθογόνων διεργασιών ή φαρμακολογικών απαντήσεων σε θεραπευτική παρέμβαση» (NIH, 2001) . Οι καρκινικοί

βιοδείκτες υποστηρίζουν τις αποφάσεις που σχετίζονται με τη διάγνωση και τη θεραπεία με τρεις κύριους σκοπούς, καθορίζοντας: ποιους θα θεραπεύσει (προγνωστικοί δείκτες), πώς θα θεραπεύσει (προγνωστικοί δείκτες), και πόση θεραπεία θα χρειαστεί (φαρμακοδυναμικοί δείκτες).



Διάγραμμα 2.β Κατηγορίες βιοδεικτών

Οι βιοδείκτες χρησιμοποιούνται από όλους τους ιατρικούς και παραϊατρικούς κλάδους για να απαντήσουν σε τρεις βασικές ερωτήσεις. **Πόσο, τι και για ποιόν.**

Πολλοί από τους βιοδείκτες είναι εύκολο να μετρηθούν και αποτελούν μέρος ιατρικών εξετάσεων ρουτίνας. Οι βιοδείκτες συνδυαστικά ή μεμονωμένα μπορούν να παρέχουν ένα σύστημα έγκαιρης προειδοποίησης για την εμφάνιση ασθενειών σε πληθυσμιακές ομάδες με συγκεκριμένα χαρακτηριστικά. Η πίεση του αίματος, ο καρδιακός ρυθμός και ο παλμός είναι δείκτες της καρδιαγγειακής λειτουργίας. Η χοληστερόλη και τα τριγλυκερίδια παρέχουν πληροφορίες σχετικά με τις μεταβολικές διεργασίες και χρησιμοποιούνται για την αξιολόγηση του κινδύνου στεφανιαίας νόσου. Οι μετρήσεις του σώματος όπως το βάρος, ο δείκτης μάζας σώματος (BMI) και ο λόγος μέσης προς ισχίου χρησιμεύουν ως δείκτες παχυσαρκίας, χρόνιων μεταβολικών διαταραχών και λιπαρών καταθέσεων. Οι μετρήσεις των T-κυττάρων (T-cell) παρέχουν πληροφορίες σχετικά με την κατάσταση του ανοσοποιητικού συστήματος. η κορτιζόλη, μια στεροειδής ορμόνη, συχνά παράγεται ως απάντηση στο χρόνιο στρες. (Today's Research on Aging 2008)

Ένας μεμονωμένος βιοδείκτης, αφού ξεπεράσει ένα συγκεκριμένο όριο (threshold) αποτελεί ένδειξη κινδύνου για μελλοντική ασθένεια. Με την συνεκτίμηση και άλλων δεικτών αλλά και παραμέτρων οι οποίοι συλλέγονται στα πλαίσια ενός ατομικού ιστορικού, μπορεί να δημιουργηθεί ένας νέος δείκτης που δίνει πιο συγκεκριμένες πληροφορίες για την υγεία ενός ατόμου. Για παράδειγμα από τις μελέτες προκύπτει ότι οι μεγαλύτεροι σε ηλικία άντρες έχουν υψηλότερο κίνδυνο εμφάνισης σοβαρών επιπτώσεων στην υγεία τους λόγω συνδυασμού διαταραχών στη λειτουργία του ανοσοποιητικού συστήματος και του νευροενδοκρινικού συστήματος. Σε μεγαλύτερες γυναίκες, οι υψηλοί κίνδυνοι υγείας οφείλονται συνήθως στην υψηλή συστολική αρτηριακή πίεση σε συνδυασμό με τις βλάβες που σηματοδοτούνται από άλλους βιοδείκτες.

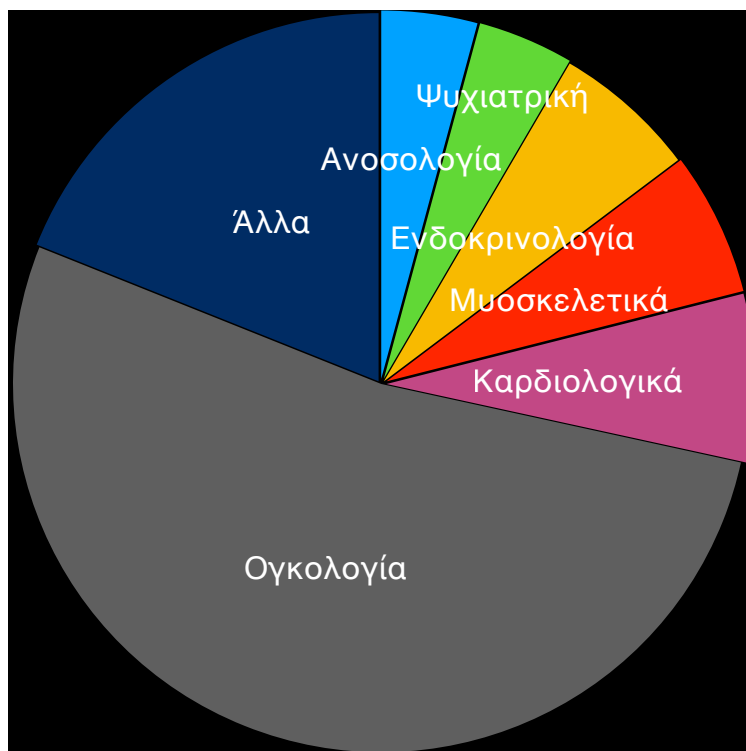
ΒΙΟΔΕΙΚΤΕΣ	ΠΕΡΙΓΡΑΦΗ	ΑΠΟΤΕΛΕΣΜΑΤΑ ΥΓΕΙΑΣ
Συστολική αρτηριακή πίεση(SBP)	Δείκτης καρδιαγγειακής δραστηριότητας: μέγιστη πίεση σε μια αρτηρία όταν η καρδιά τροφοδοτεί το σώμα με αίμα	Καρδιαγγειακο θάνατο(CVD), εγκεφαλικό επεισόδιο, στεφανιαία καρδιακή νόσο(CHD)
Διαστολική αρτηριακή πίεση	Δείκτης καρδιαγγειακής δραστηριότητας: μέγιστη πίεση σε μια αρτηρία όταν η καρδιά αναπαύεται	Καρδιαγγειακο θάνατο(CVD), εγκεφαλικό επεισόδιο, στεφανιαία καρδιακή νόσο(CHD)
Ρυθμός παλμών ανάπαυσης	δείκτης της λειτουργίας της καρδιάς και μέτρηση της συνολικής σωματικής ικανότητας	Στεφανιαία καρδιακή νόσο
Συνολική χοληστερόλη	Βοηθά στη σύνθεση χολικών οξέων και στεροειδών ορμονών	Στη μέση ηλικία: chd και θνησιμότητα όλων των αιτιών στη μεσαία ηλικία: chd και θνησιμότητα όλων των αιτιών σε μεγαλύτερη ηλικία: σχήμα u σχέσης με το θάνατο
Χαμηλής πυκνότητας λιποπρωτεΐνη(LDL)	Μεταφέρει τη χοληστερόλη από το ήπαρ ώστε να ενσωματωθεί στους ιστούς κυτταρικής μεμβράνης	Chd, αθηροσκλήρυνση, εγκεφαλικό επεισόδιο, περιφερική αγγειακή νόσο
Υψηλής πυκνότητας λιποπρωτεΐνη(HDL)	προστατευτική χοληστερόλη προστατευτική χοληστερόλη	Χαμηλότερο αθηροσκληρωτικό cvd

ΒΙΟΔΕΙΚΤΕΣ	ΠΕΡΙΓΡΑΦΗ	ΑΠΟΤΕΛΕΣΜΑΤΑ ΥΓΕΙΑΣ
Τριγλυκερίδια	λιπαρές ουσίες που αποθηκεύονται για χρήση ενέργειας λιπαρές ουσίες που αποθηκεύονται για χρήση ενέργειας	Καρδιακή προσβολή, chd, ασθένεια στεφανιαίας αρτηρίας, παγκρεατίτιδα
Γλυκόζη νηστείας	μετρά την ποσότητα της ζάχαρης στον δείκτη του διαβήτη μετρά την ποσότητα της ζάχαρης στον δείκτη του διαβήτη	Διαβήτη, chd, θνησιμότητα, κακή γνωστική λειτουργία
Δείκτης μάζας σώματος(BMI)	Δείκτη ισορροπίας μεταξύ της πρόσληψης ενέργειας και της ενεργειακής δαπάνης	CVD, σακχαρώδη διαβήτη, εγκεφαλικό επεισόδιο, θνησιμότητα, κάποιους καρκίνους, οστεοαρθρίτιδα
Αναλογία μέσης ισχίου	δείκτη κοιλιακής παχυσαρκίας δείκτη κοιλιακής παχυσαρκίας	Υπέρταση, chd, διαβήτη που εξαρτάται από τη έλλειψη ινσουλίνη και εγκεφαλικό επεισόδιο
T-Cells	λευκά αιμοσφαίρια που προστατεύουν από παθογόνους και όγκους	Καρκίνο, θάνατο, αρτηριοσκλήρωση, Αλτσχάιμερ
Κορτιζόλη	στεροειδής ορμόνη που αντανάκλα την ανταπόκριση του οργανισμού στο φυσιολογικό στρες	CVD, κακή γνωστική λειτουργία, κατάγματα, λειτουργική αναπηρία, θάνατο

ΒΙΟΔΕΙΚΤΕΣ	ΠΕΡΙΓΡΑΦΗ	ΑΠΟΤΕΛΕΣΜΑΤΑ ΥΓΕΙΑΣ
Ηλεκτροκαρδιογράφημα(EKG)	μέτρηση των ηλεκτρικών παλμών στην καρδιά μέτρηση των ηλεκτρικών παλμών στην καρδιά	καρδιαγγειακό κίνδυνο, εγκεφαλικό επεισόδιο, θνησιμότητα

Πίνακας 2.γ Βιοδείκτες σε εξετάσεις ρουτίνας

Οι βιοδείκτες χρησιμοποιούνται συστηματικά στις περισσότερες ιατρικές ειδικότητες. Στην ογκολογία όμως έχουν μπει πλέον σε συστηματική κλινική χρήση και αποτελούν σημαντικό εργαλείο διαχείρισης ασθενών.



Διάγραμμα 2.δ Βιοδείκτες και ιατρικοί κλάδοι

Όπως φαίνεται και στο **διάγραμμα 2.δ** περισσότερο από το 50% των χρησιμοποιούμενων βιοδεικτών έχουν πρακτική εφαρμογή στον κλάδο της ογκολογίας. (www.clinicaltrials.gov). Περισσότερο από το 1/3 των κλινικών μελετών για τον καρκίνο περιλαμβάνουν βιοδείκτες.

Τα γενετικά τεστ για την πρόβλεψη της ευαισθησίας ενός ατόμου σε μια συγκεκριμένη ασθένεια ή για τον εντοπισμό συγκεκριμένων γενετικών μεταλλάξεων αποτελεί μία από τις πιο παραγωγικές μεθόδους της ιατρικής. Επίσης, οι γενετικές εξετάσεις και η εικονική μοντελοποίηση ασθενειών μπορούν να προβλέψουν τη αντίδραση ενός ασθενούς σε ορισμένα φάρμακα και να οδηγήσουν σε ενίσχυση της θεραπευτικής αγωγής και της αποτελεσματικότερης διαχείρισης του προβλήματος.

Μια πραγματικά εξατομικευμένη προσέγγιση στη γονιδιωματική ιατρική θα ενισχύσει σημαντικά την ποιότητα ζωής των καρκινοπαθών. Παραδείγματα εξατομικευμένων αγωγών καρκίνου περιλαμβάνουν βιοδείκτες για τον εντοπισμό καρκίνου του μαστού οι οποίοι υποβάλλονται σε αγωγή με συγκεκριμένο φάρμακο (αντίσωμα trastuzumab (Herceptin™)) (Martinez & Rossbach 2012).

Προκλήσεις στην ευρεία χρήση βιοδεικτών υπάρχουν και είναι σημαντικές. Βασικές παράμετροι επίσης για την καλύτερη κατανόηση και καλύτερη εφαρμογή των βιοδεικτών και στον συνδυαστικό προσδιορισμό πληθυσμιακών κριτηρίων παίζει η ηλικία, η φυλή, το φύλο και άλλοι εξωγενείς παράγοντες όπως το επίπεδο και ο τρόπος ζωής, το βιοτικό επίπεδο. (Geronimus et.al. 2007). Για τη δημιουργία εφαρμογών

γενικής χρήσης και τη διαμόρφωση μιας πιο συστηματικής προσέγγισης θα πρέπει όλες οι παράμετροι να αξιολογηθούν και να διερευνηθούν.

Στην Αμερική για παράδειγμα οι μαύροι σε σχέση με τους λευκούς είναι πιθανότερο να αντιμετωπίσουν στρεσογόνες καταστάσεις όπως φυλετικές διακρίσεις στην αγορά εργασίας και στην εύρεση στέγης, περιθωριοποίηση. Η κακή ψυχολογία και το στρες θα έχει βιολογική επίπτωση με ότι αυτό συνεπάγεται. Η ηλικία επίσης αυξάνει το ρίσκο εμφάνισης κάποιας αναπηρίας ή χρόνιας ασθένειας. Υπάρχουν ενδείξεις ότι περιβαλλοντολογικές και ψυχολογικές αιτίες η συσσώρευση των οποίων σε βάθος χρόνου μπορεί να επηρεάσει αρνητικά ένα βιολογικό σύστημα.

2.2 Τεχνολογία sequencing επόμενης γενιάς

Σήμερα, μέσω ανάλυσης του γονιδιώματος (NGS-next generation sequencing) μπορούν να αναλυθούν εκατοντάδες εκατομμύρια θραυσμάτων DNA ανά κύκλο. Οι τεχνολογίες επόμενης γενιάς στην ανάλυση ακολουθίας του ανθρώπινου γονιδιώματος μπορούν να παράξουν έως 100 φορές περισσότερα δεδομένα σε σχέση με τις προηγούμενες μεθόδους αναλύσεις (π.χ μέθοδος Sanger). Αυτό επέτρεψε στους ερευνητές να αλληλοσυμπληρώσουν μεμονωμένα γονιδιώματα και να εντοπίσουν συνδυασμούς μεταλλάξεων με συγκεκριμένες ασθένειες. Καθώς ο καρκίνος είναι εγγενώς μια ασθένεια του γονιδιώματος, δεν αποτελεί έκπληξη το γεγονός ότι η τεχνολογία NGS εφαρμόζεται ήδη στην έρευνα για τον καρκίνο με υποσχέσεις μεγαλύτερης κατανόησης της καρκινογένεσης. Τα διαγνωστικά με βάση την ανάλυση του γονιδιώματος έχουν τεράστια προοπτική στη βελτίωση της διαχείρισης του καρκίνου (Perkins et.al. 2017).

Η τεχνολογία ακολουθίας επόμενης γενιάς (NGS) βασίζεται στην κοπή DNA σε μικρά θραύσματα και στη μαζική παράλληλη αλληλούχισή τους. Τα πολλαπλά επικαλυπτόμενα τμήματα που ονομάζονται "αναγνώσματα" συναρμολογούνται σε συνεχή ακολουθία. Για να μειωθούν τα σφάλματα προσδιορισμού της αλληλουχίας, κάθε περιοχή γονιδιώματος πρέπει να αλληλουχηθεί αρκετές δεκάδες φορές. Αυτή η προσέγγιση αλληλουχίας βασίζεται στην υπόθεση ότι τα θραύσματα γονιδιωματικού DNA είναι τυχαία και ανεξάρτητα από την αλληλουχία.

Εν συντομία, η τεχνολογία NGS μπορεί να χωριστεί σε τρεις φάσεις: προετοιμασία προτύπου, ανάλυση αλληλουχίας και ανίχνευση και ανάλυση αλληλουχίας. Για την προετοιμασία του προτύπου, είναι απαραίτητη η παραγωγή αντιπροσωπευτικού, μη αντικειμενικού συνόλου κατακερματισμένου DNA. Για να επιτραπεί η παράλληλη επεξεργασία μεγάλου αριθμού αντιδράσεων προσδιορισμού αλληλουχίας, αυτά τα θραύσματα ακινητοποιούνται ή περιορίζονται σε χωρικά διαχωρισμένες θέσεις σε στερεό υπόστρωμα ή σε θάλαμο νανοσωματιδίων. Τα χωρικά διαχωρισμένα πρότυπα στη συνέχεια είτε ενισχύονται κλωνικά είτε παραμένουν ως απλά μόρια. Οι εμπορικά διαθέσιμες πλατφόρμες χρησιμοποιούν μια ποικιλία αρχών για την πραγματική αντίδραση αλληλούχησης (Porotsova 2014). Αυτά περιλαμβάνουν τον κυκλικό αναστρέψιμο τερματισμό, την προσθήκη ενός νουκλεοτιδίου, την αλληλούχισή σε πραγματικό χρόνο και την αλληλούχισή με σύνδεση.

Το γονιδίωμα ενός ζωντανού οργανισμού μοιάζει με βιβλιοθήκη γεμάτη με βιβλία, τα χρωμοσώματα που περιέχουν κείμενα που αποτελούνται από γράμματα και ονομάζονται νουκλεοτίδια. Οι πρώτες μέθοδοι αποκρυπτογράφησης βιολογικών αλληλουχιών βασίστηκαν στην ακριβή εκτομή ενός συγκεκριμένου θραύσματος DNA και στην ακριβή ανάγνωσή του. Μια εναλλακτική και αρχικά φαινομενικά ασυνεπής μέθοδος προσδιορισμού αλληλουχίας προτάθηκε το 1979, όπου τα πολλαπλά αντίγραφα ενός DNA ολόκληρου γονιδιώματος θα διασπαστούν σε μικρά θραύσματα, τα οποία τοποθετούνται σε αλληλουχία και στη συνέχεια αυτές οι αλληλουχίες (που

ονομάζονται "αναγνώσματα") συναρμολογούνται σε ένα συνεχές κείμενο με βάση τα επικαλυπτόμενα άκρα. Παρ' όλα αυτά, η εκρηκτική ανάπτυξη αυτοματοποιημένων αλληλουχιών και η πρόοδος στον κλάδο της πληροφορικής καθόρισαν την σημερινή κυριαρχία αυτής της μεθόδου, που ονομάζεται τυχαία ακολουθία (shotgun sequencing). Οι σύγχρονες μηχανές προσδιορισμού αλληλουχίας είναι σε θέση να διαβάζουν εκατοντάδες εκατομμύρια αναγνώσματα (reads) ανά ημέρα, όπου κάθε ανάγνωση αποτελείται από δεκάδες ή εκατοντάδες νουκλεοτίδια.(non random DNA fragmentation).

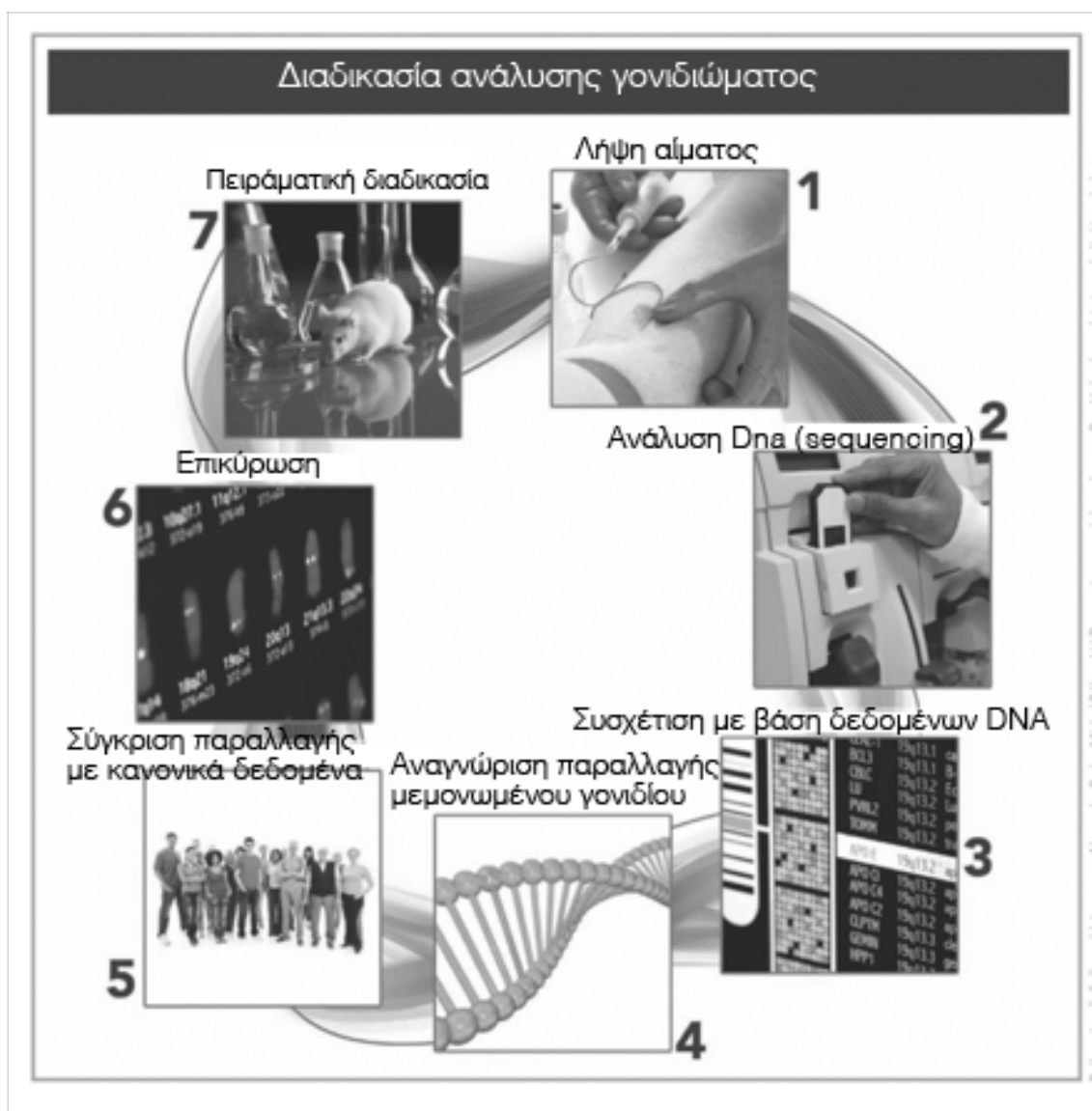
Υπάρχουν δύο κύριοι τύποι εργαστηριακών προσεγγίσεων όπως είδαμε και παραπάνω, η αλληλουχία ολόκληρου του γονιδιώματος DNA - και η στοχευμένη αλληλουχία. Η αλληλουχία ολόκληρου του γονιδιώματος, όπως υποδηλώνει και το όνομά, παρέχει έναν ολοκληρωμένο χαρακτηρισμό ολόκληρου του γονιδιώματος. Αυτό επιτρέπει μεγαλύτερη ικανότητα ταυτοποίησης αλλοιώσεων τόσο σε περιοχές κωδικοποίησης όσο και σε μη κωδικοποιημένες περιοχές. Η στοχευμένη αλληλουχία (shotgun sequencing) είναι μια προσέγγιση χαμηλότερου κόστους η οποία βασίζεται στην αλληλουχία τυχαία παραγόμενων θραυσμάτων. Οι ερευνητές επιλέγουν συχνά την αλληλουχία αυτή λόγω της απλούστερης προετοιμασίας της βιβλιοθήκης αλληλουχίας, του χαμηλότερου κόστους και της μειωμένης ποσότητας του απαιτούμενου DNA εισόδου σε σύγκριση με την παραδοσιακή ακολουθία ολόκληρου του γονιδιώματος αλλά είναι λιγότερο ακριβής σε σχέση με την αλληλουχία ολόκληρου του γονιδιώματος (Y Kim et.al 2011).

Οι τεχνολογικές εξελίξεις της NGS οδηγούν επίσης στην ανάπτυξη νέων διαγνωστικών και θεραπευτικών προσεγγίσεων για τη θεραπεία του καρκίνου (American laboratory 2015), καθώς οι ερευνητές ανασυντάσσουν τη σύγκριση όγκων και φυσιολογικών γονιδιωμάτων με συγκεκριμένους τύπους καρκίνου (Pfeifer and Hainaut 2011). Στο γονιδίωμα του καρκίνου, υπάρχει η δυνατότητα για ταχεία ταυτοποίηση των αλλαγών σε συγκεκριμένους ασθενείς με καρκίνο στους συμπαγείς όγκους (Bubnoff 2008). Μια άλλη ενδιαφέρουσα εφαρμογή της τεχνολογίας είναι οι «εξατομικευμένοι» βιοδείκτες, οι οποίοι αναπτύχθηκαν πρόσφατα για να ανιχνεύσουν την παρουσία ειδικών για τον όγκο γονιδιακών αναδιατάξεων σε δείγματα πλάσματος από ασθενείς. Οι γονιδιωματικές αναδιατάξεις εντοπίστηκαν ειδικά στον όγκο, οι οποίες δεν φαινόταν στον κανονικό σωματικό ιστό του ασθενούς. Οι ψηφιακές δοκιμασίες PCR σχεδιάστηκαν στη συνέχεια σε σημεία διακοπής της αναδιάταξης για να παράσχουν έναν ευαίσθητο καρκινικό βιοδείκτη ο οποίος χρησιμοποιήθηκε επιτυχώς για την παρακολούθηση της υπολειμματικής νόσου μετά από ολοκλήρωση θεραπείας (Deborah Levenson 2012). Επιπλέον, η τεχνολογία HT-NGS επιτρέπει επίσης τον εντοπισμό των αιτιωδών μεταλλάξεων που είναι υπεύθυνες για την πρόκληση της μετάστασης του καρκίνου και δημιουργεί σημαντικές προσδοκίες για τη βελτίωση των ογκολογικών αποτελεσμάτων.

Το ριβονουκλεϊκό οξύ (RNA) είναι ένα πολυμερές μόριο απαραίτητο σε διάφορους βιολογικούς ρόλους στην κωδικοποίηση, αποκωδικοποίηση, ρύθμιση και έκφραση γονιδίων. Το RNA και το DNA είναι νουκλεϊκά οξέα και, μαζί με τα λιπίδια, οι πρωτεΐνες και οι υδατάνθρακες, αποτελούν τα τέσσερα κύρια μακρομόρια απαραίτητα για όλες τις γνωστές μορφές ζωής. Όπως το DNA, το RNA συναρμολογείται ως μια αλυσίδα νουκλεοτιδίων, αλλά σε αντίθεση με το DNA, απαντάται συχνότερα στη φύση ως ένας μονόκλωνος διπλωμένος στον εαυτό του, παρά σε έναν ζευγαρωμένο διπλό κλώνο. Οι κυτταρικοί οργανισμοί χρησιμοποιούν αγγελιαφόρο RNA (mRNA) για τη μετάδοση γενετικών πληροφοριών (χρησιμοποιώντας τις αζωτούχες βάσεις της γουανίνης, της ουρακίλης, της αδενίνης και της κυτοσίνης, που υποδηλώνεται με τα γράμματα G, U, A και C) που κατευθύνει τη σύνθεση συγκεκριμένων πρωτεϊνών. Πολλοί ιοί κωδικοποιούν τη γενετική τους πληροφορία χρησιμοποιώντας ένα γονιδίωμα RNA. Ανάλογα με το μήκος της αλυσίδας RNA, το RNA περιλαμβάνει μικρό RNA και μακρύ RNA.

Το μικρο RNA θα μπορούσε να παράσχει μια περαιτέρω εφαρμογή του NGS στην ανακάλυψη νέων βιοδεικτών. Τα μικρο-RNAs (miRNAs) εμπλέκονται στην έλεγχο της πρωτεϊνικής μετάφρασης και υπάρχουν στο πλάσμα του αίματος. Το NGS θα μπορούσε να χρησιμοποιηθεί για την ανίχνευση ιστών ή πλάσματος αίματος για τη δημιουργία προφίλ miRNA ολόκληρου του γονιδίου, το οποίο στη συνέχεια θα μπορούσε να εξορύσσεται για συγκεκριμένες υπογραφές βιολογικού σήματος.

Στην **εικόνα 2.5** φαίνεται συνοπτικά η διαδικασία ανάλυσης του γονιδιώματος μέσω λήψης αίματος. Μετά τη λήψη γίνεται η ανάλυση σε ειδικά μηχανήματα όπως έγινε η περιγραφή παραπάνω. Συσχετίζονται τα αποτελέσματα με την υπάρχουσα βάση δεδομένων ώστε να ταυτοποιηθούν τα σχετικά ευρήματα, να αναγνωριστούν οι μεταλλάξεις και να επικυρωθούν (National Human Genome Research Institute 2016).



Εικόνα 2.5 Χαρτογράφηση διαδικασιών ανάλυσης γονιδιώματος

2.3 Παχυσαρκία

Η παχυσαρκία αντιπροσωπεύει μείζον πρόβλημα υγείας, όχι μόνο στη Δύση, αλλά όλο και περισσότερο στις χώρες με χαμηλό και μεσαίο εισόδημα. Προκειμένου να αναπτυχθούν επιτυχημένες στρατηγικές για την απώλεια βάρους, είναι σημαντικό να κατανοηθεί η μοριακή παθογένεση της αλλαγής βάρους. Ορισμένα μονοπάτια, που συνεπάγονται οξειδωτικό στρες αλλά και το θεμελιώδες ρυθμιστικό πλαίσιο της ινσουλίνης, έχουν εμπλακεί στην αύξηση βάρους και στη ρύθμιση των ενεργειακών δαπανών. Επιπροσθέτως, ένα σημαντικό σώμα εργασίας έχει επισημάνει το ρόλο των μεταβολιτών που παράγονται από το μικρόβιο του εντέρου, ιδιαίτερα των λιπαρών οξέων βραχείας αλυσίδας, και στις δύο διαδικασίες. Η τρέχουσα ανασκόπηση παρέχει μια σύντομη κατανόηση των μηχανισμών στις οποίες βασίζονται οι συσχετισμοί της αλλαγής βάρους με αλλαγές στον μεταβολισμό των λιπιδίων και των αμινοξέων, του μεταβολισμού της ενέργειας, της διατροφικής σύνθεσης και της δυναμικής της ινσουλίνης, καθώς και την επίδραση του μικροβίου του εντέρου.

Η παχυσαρκία, που αποτελεί σημαντικό πρόβλημα δημόσιας υγείας, οφείλεται σε υψηλότερη πρόσληψη ενέργειας από την σχετική δαπάνη της για μεγάλο χρονικό διάστημα. Παρόλο που πολλά άτομα σε βιομηχανικά έθνη είναι υπέρβαρα ή παχύσαρκα, ένα σημαντικό ποσοστό ατόμων με φυσιολογικό βάρος δεν γίνονται ποτέ υπέρβαρα, αντανακλώντας εν μέρει τη μεγάλη διακύμανση μεταξύ των ατόμων με πλεόνασμα θερμιδικής πρόσληψης. Η απώλεια βάρους και η αύξηση βάρους συνδέονται με μειώσεις και αυξήσεις στις ενεργειακές δαπάνες και πρόσληψης. Ωστόσο, πολλοί διαφορετικοί παράγοντες συμβάλλουν στην ομοιόσταση του σωματικού βάρους στους ανθρώπους και, καθώς αναπτύσσεται η παχυσαρκία, συμβαίνουν ορισμένες μεταβολικές αλλαγές, οι οποίες ενδέχεται να μην αντιστραφούν εντελώς όταν χάνεται βάρος (Vijay 2019).

Το μεταβολικό σύνδρομο αντιπροσωπεύει σημαντικό αριθμό θανάτων και ασθενειών στις δυτικές χώρες και επίσης, όλο και περισσότερο, σε χώρες με χαμηλότερα εισοδήματα. Αυτές οι χώρες, όπως η Ινδία, αντιμετωπίζουν διπλή επιβάρυνση, με μεγάλο αριθμό υποσιτισμένων ατόμων σε αγροτικές περιοχές και αυξανόμενο αριθμό ατόμων που πλήττονται από παχυσαρκία και ασθένειες που σχετίζονται με την παχυσαρκία.

Σχεδόν το 40% των ενηλίκων έχει δείκτη μάζας σώματος που τους χαρακτηρίζει υπέρβαρους και 13% ως πάσχοντες από παχυσαρκία. Δεδομένης της αύξησης του κινδύνου σακχαρώδους διαβήτη, οστεοαρθρίτιδας και καρδιαγγειακών παθήσεων που προκαλείται από την παχυσαρκία, υπάρχει ανάγκη κατανόησης των μοριακών προσδιοριστικών παραγόντων της αλλαγής βάρους.

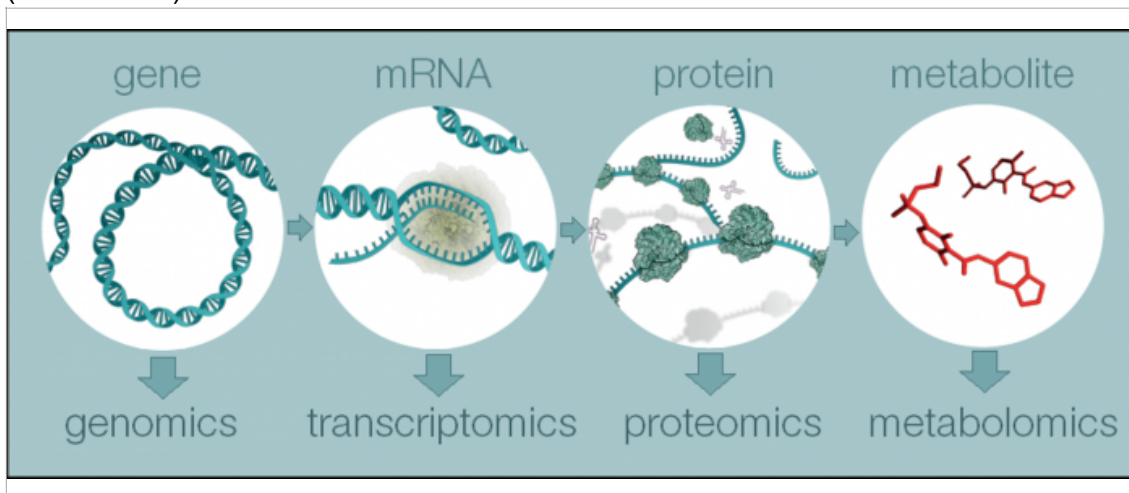
2.4 διατροφικοί βιοδείκτες

2.4α Metabolomics -μεταβολικό προφίλ

Το μεταβολικό προφίλ αναφέρεται στην συστηματική μελέτη των μοναδικών χημικών δακτυλικών αποτυπωμάτων γνωστών ως μεταβολιτών που αφήνουν πίσω συγκεκριμένες κυτταρικές διεργασίες. Το μεταβολικό προφίλ μπορεί να δώσει ένα στιγμιότυπο της φυσιολογίας του κυττάρου, και κατ'επέκταση μια άμεση «λειτουργική ανάγνωση της φυσιολογικής κατάστασης» ενός οργανισμού.

Ο προσδιορισμός του μεταβολικού προφίλ αντανακλά τις μεταβολικές διεργασίες, συμπεριλαμβανομένων των αλλαγών που εμπλέκονται στην παθολογία. Ορισμένες επιστημονικές δημοσιεύσεις που χαρακτηρίζουν τους μεταβολίτες μέχρι

σήμερα έχουν επικεντρωθεί στην αλληλοεπικάλυψη μεταξύ σακχαρώδους διαβήτη τύπου 2 και παχυσαρκίας. Αρκετοί μεταβολίτες εμπίπτουν σε αυτήν την κατηγορία, όπως αμινοξέα διακλαδισμένης αλυσίδας (BCAAs), γλουταμίνη, προλίνη, κυστεΐνη, τυροσίνη, θρεονίνη, φαινυλαλανίνη, τρυπτοφάνη, παντοθενικό οξύ και χολίνη, τα οποία αυξάνονται τόσο στην παχυσαρκία όσο και στον διαβήτη, ενώ η γλυκίνη, η ασπαραγίνη, η κιτροουλίνη και η μεθειονίνη μειώνονται στον διαβήτη και την παχυσαρκία. Σε μια εγκάρσια μελέτη 947 συμμετεχόντων, 37 μεταβολίτες συσχετίστηκαν σημαντικά με τον δείκτη μάζας σώματος, συμπεριλαμβανομένων δεκαεννέα λιπιδίων, δώδεκα αμινοξέων και έξι άλλων. Δεκαοκτώ από αυτούς τους συσχετισμούς δεν είχαν αναφερθεί προηγουμένως, συμπεριλαμβανομένων της ιστιδίνης και της βουτυρυλκαρνιτίνης. (Moore 2014)



Εικόνα 2.ε Από τα genomics στα metabolomics

Ακριβώς όπως η γονιδιωματική (genomics) είναι η μελέτη του DNA και των γενετικών πληροφοριών σε ένα κύτταρο, και η μεταγραφική (transcriptomics) είναι η μελέτη του RNA και των διαφορών στην έκφραση mRNA, η μεταβολική (metabolomic) είναι η μελέτη υποστρωμάτων και προϊόντων μεταβολισμού, τα οποία επηρεάζονται τόσο από γενετικούς όσο και από περιβαλλοντικούς παράγοντες. Τα Metabolomics είναι μια ισχυρή προσέγγιση επειδή οι μεταβολίτες και οι συγκεντρώσεις τους, σε αντίθεση με άλλα μέτρα «omics», αντανakλούν άμεσα την υποκείμενη βιοχημική δραστηριότητα και την κατάσταση των κυττάρων / ιστών. Έτσι, η μεταβολική αντιπροσωπεύει καλύτερα τον μοριακό φαινότυπο.

Η μεταβολική περιλαμβάνει αναλυτικές τεχνολογίες που μπορούν να παρέχουν διαγνωστικά πρότυπα μέσω δακτυλικών αποτυπωμάτων, απόλυτου ποσοτικού προσδιορισμού στοχευμένων μεταβολιτών μέσω ανάλυσης συγκέντρωσης, σχετικού ποσοτικού προσδιορισμού μεγάλων μερίδων του μεταβολισμού χρησιμοποιώντας προφίλ μεταβολιτών και ιχνηλάτηση μεμονωμένων μεταβολιτών μέσω μεταβολικού συστήματος ανάλυσης ροής. Κάθε μία από αυτές τις τεχνολογίες υποστηρίζεται από τις δύο πιο συχνά χρησιμοποιούμενες και ισχυρές τεχνικές που διατίθενται σήμερα: τη φασματομετρία μάζας και το NMR.

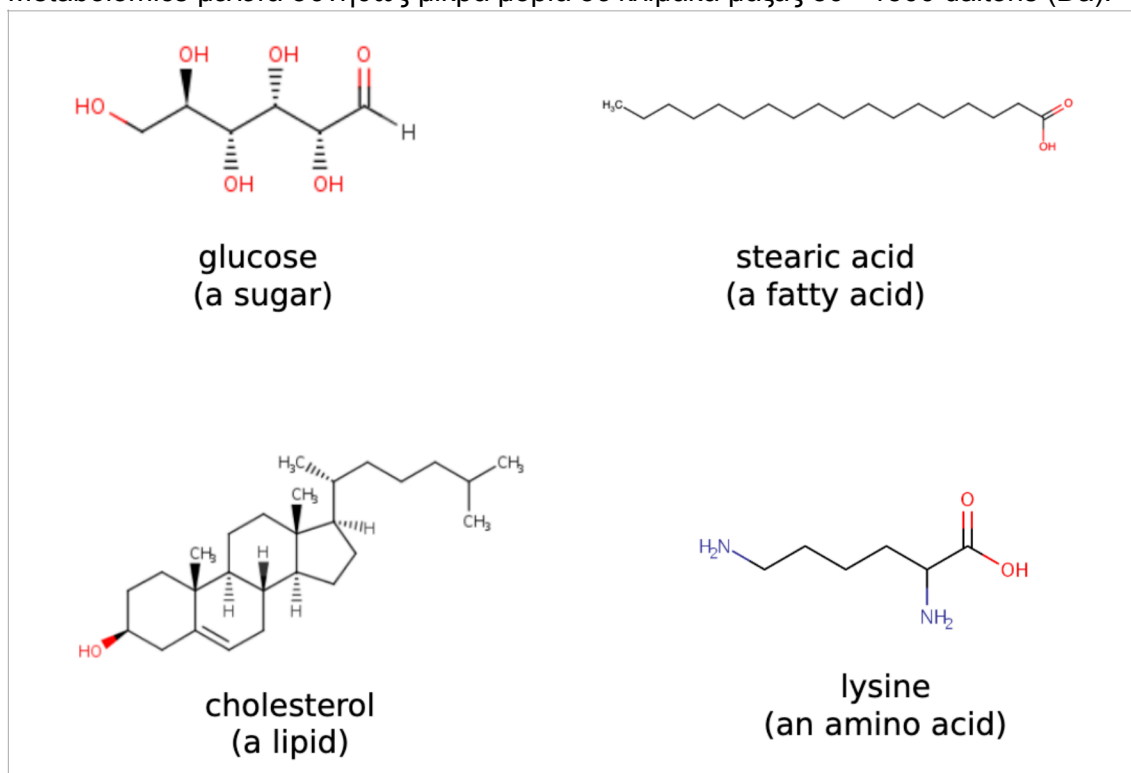
Οι τεχνικές μεταβολισμού με βάση τη φασματομετρία μάζας είναι οι πιο ευαίσθητες για ταυτόχρονη ανάλυση μεγάλου αριθμού ενώσεων. Ενώ περιορίζεται στις δυνατότητες ποσοτικοποίησης χωρίς κατάλληλα επισημασμένα πρότυπα, η τεράστια ποσότητα πληροφοριών που διατίθεται σε ένα μόνο πείραμα LC-MS ή GC-MS μπορεί να παρέχει λεπτομερείς πληροφορίες σχετικά με τα πρότυπα αλλαγής μεταβολίτη σε ένα ολόκληρο μεταβολικό δίκτυο.

Ο μεταβολισμός NMR συμπληρώνει τη φασματομετρία μάζας. Είναι περιορισμένη μέθοδος ως προς την ευαισθησία, αλλά είναι μοναδικά ικανή να αποσαφηνίσει τη μοριακή δομή. Ένα σημαντικό πρόσθετο χαρακτηριστικό της NMR

είναι ότι είναι ποσοτικά ικανή να παρέχει απόλυτα επίπεδα ανιχνευόμενων ενώσεων όταν χρησιμοποιούνται κατάλληλες τεχνικές.

2.4β Metabolites - μεταβολίτες

Ένα μικρό μόριο (ή μεταβολίτης) είναι μια οργανική ένωση χαμηλού μοριακού βάρους, που συνήθως εμπλέκεται σε μια βιολογική διεργασία ως υπόστρωμα ή προϊόν. Η Metabolomics μελετά συνήθως μικρά μόρια σε κλίμακα μάζας 50 - 1500 daltons (Da).



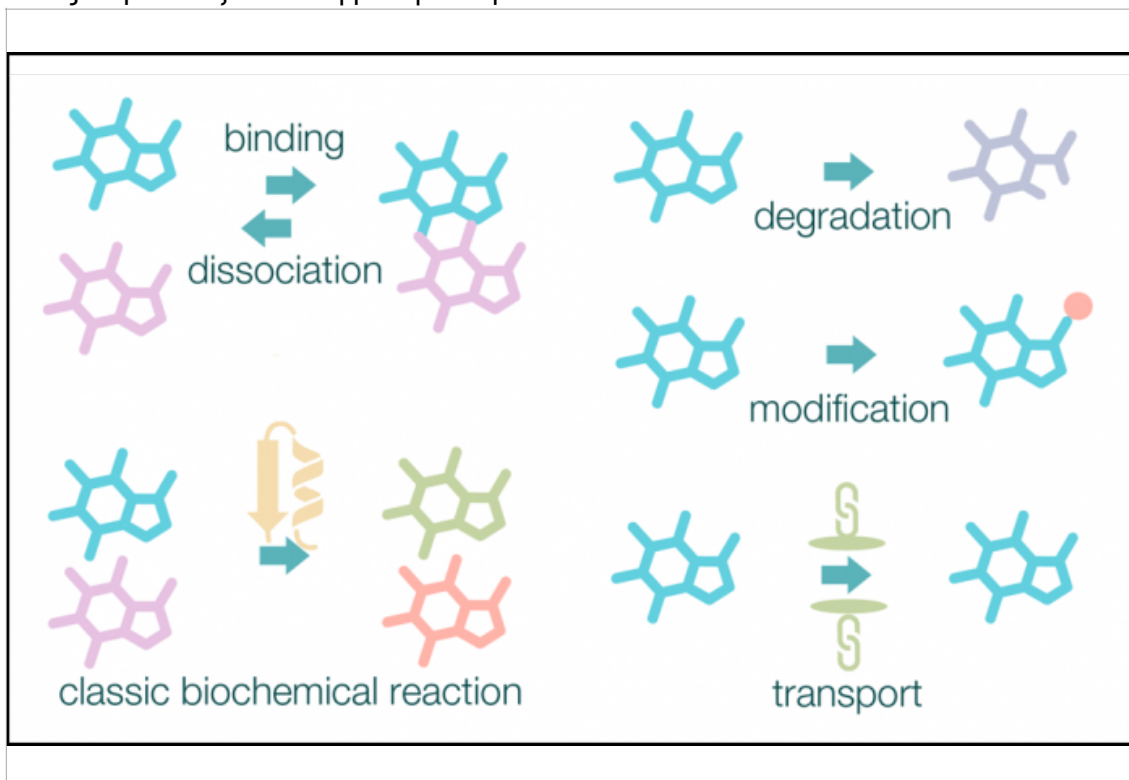
Εικόνα 2.ζ Metabolites

Παραδείγματα μικρών μορίων φαίνονται στο παραπάνω σχήμα και περιλαμβάνουν:

- A. σάκχαρα
- B. λιπίδια
- C. αμινοξέα
- D. φαινολικές ενώσεις
- E. αλκαλοειδή

Υπάρχει μεγάλη διακύμανση στους μεταβολίτες μεταξύ των ειδών, εκτιμάται ότι υπάρχουν περίπου 200.000 μεταβολίτες σε ολόκληρο το φυτικό βασίλειο και κάπου μεταξύ 7.000 και 15.000 σε ένα μεμονωμένο φυτικό είδος. Αντίθετα, στους ανθρώπους, πιστεύεται ότι υπάρχουν περίπου 3.000 ενδογενείς ή κοινοί μεταβολίτες. Αυτές οι εκτιμήσεις είναι προσεγγίσεις που είναι πιθανό να υποτιμώνται επειδή είναι δύσκολο να ανιχνευθούν μόρια χαμηλής αφθονίας. Ωστόσο, μπορεί να συναχθεί το συμπέρασμα

ότι τα φυτά είναι ιδιαίτερα βιοχημικά πλούσια σε σύγκριση με πολλά άλλα είδη. Περιέχουν επίσης συνήθως μεγαλύτερο αριθμό γονιδίων από άλλα ευκαρυωτικά. Ο μεταβολισμός είναι το πλήρες σύνολο μεταβολιτών εντός κυττάρου, ιστού ή βιολογικού δείγματος σε οποιοδήποτε δεδομένο χρονικό σημείο. Ο μεταβολισμός είναι εγγενώς πολύ δυναμικός: μικρά μόρια απορροφώνται συνεχώς, συντίθενται, αποικοδομούνται και αλληλεπιδρούν με άλλα μόρια, τόσο εντός όσο και μεταξύ βιολογικών συστημάτων, και με το περιβάλλον. Οι κύριες μεταβολικές αντιδράσεις απεικονίζονται στο σχήμα παρακάτω. Αυτές οι κυτταρικές αντιδράσεις εμφανίζονται όπως παρουσιάζονται στη βάση δεδομένων Reactome.



Εικόνα 2.ζ Μεταβολικές αντιδράσεις

2.4γ Nutritional genomics

Η διατροφική γονιδιωματική είναι η εφαρμογή λειτουργικών γονιδιωματικών τεχνολογιών υψηλής απόδοσης στην έρευνα για τη διατροφή. Αυτές οι τεχνολογίες μπορούν να ενσωματωθούν σε βάσεις δεδομένων γονιδιωματικών αλληλουχιών και γενετικής μεταβλητότητας μεταξύ ατόμων, επιτρέποντας στη διαδικασία της γονιδιακής έκφρασης να μελετηθεί παράλληλα για πολλές χιλιάδες διαφορετικά γονίδια. Τέτοιες τεχνικές μπορούν να διευκολύνουν τον ορισμό της βέλτιστης διατροφής σε επίπεδο πληθυσμών, συγκεκριμένων ομάδων και ατόμων. Αυτό με τη σειρά του θα προωθήσει την ανάπτυξη θεραπειών που βασίζονται στη διατροφή και τη δημιουργία βελτιωμένων τροφίμων με στόχο τη βελτίωση της υγείας.

Η διατροφή έχει σημαντικό αντίκτυπο στις χρόνιες παθήσεις και την υγεία, και η μελέτη των διατροφικών δεικτών θα μπορούσε να συνεισφέρει στον καθορισμό των βιοδραστικών συστατικών των τροφίμων. Ο ορισμός αυτών των δραστηριοτήτων θα οδηγήσει στην βελτίωση της υγείας μέσω τροποποίησης και ενίσχυσης της διατροφής.

δημιουργία νέων τροφίμων κλπ. Οι προκλήσεις εντοπίζονται στον βέλτιστο σχεδιασμό διατροφικών μελετών και στον αποτελεσματικό χειρισμό των τεράστιων συνόλων δεδομένων που δημιουργούνται, Τώρα είναι δυνατόν να οριστούν γονιδιακοί πολυμορφισμοί που προδιαθέτουν άτομα σε ασθένειες και τροποποιούν τις διατροφικές απαιτήσεις. Ο χαρακτηρισμός τέτοιων γονιδιακών πολυμορφισμών θα επιτρέψει τη στόχευση διατροφικών συμβουλών και θεραπείας σε ομάδες «σε κίνδυνο». (Elliott 2002).

2.45 Standard Biochemistry Profile - Βιοχημικό προφίλ

Στα πλαίσια αυτής της εργασίας θα χρησιμοποιηθούν βάσεις δεδομένων που έχουν σχέση με το βιοχημικό προφίλ το οποίο προκύπτει από στοχευμένες αιματολογικές εξετάσεις και προσδιορίζει τη λειτουργική ικανότητα σημαντικών οργάνων όπως το ήπαρ και τα νεφρά. Επίσης προσδιορίζει την διαφορά στον μεταβολισμό των πρωτεϊνών και του λίπους, του μεταβολισμού των υδατανθράκων, της παραγωγής ενζύμων, της ικανότητας χρήσης ενώσεων κ.λπ. Καταγράφονται παρακάτω σημαντικοί δείκτες του βιοχημικού προφίλ.

Στοιχείο	Περιγραφή
Alanine Aminotransferase	Οι μετρήσεις της αλανινοτρανσφεράσης της αλανίνης χρησιμοποιούνται στη διάγνωση και τη θεραπεία ορισμένων ηπατικών παθήσεων (π.χ. ιική ηπατίτιδα και κίρρωση) και καρδιακών παθήσεων. Τα αυξημένα επίπεδα των τρανσαμινασών μπορεί να υποδηλώνουν έμφραγμα του μυοκαρδίου, ηπατική νόσο, μυϊκή δυστροφία ή βλάβη οργάνων.
Albumin	Οι μετρήσεις της λευκωματίνης χρησιμοποιούνται στη διάγνωση και τη θεραπεία ασθενειών που εμπλέκουν το ήπαρ ή / και τα νεφρά, και χρησιμοποιούνται συχνά για την αξιολόγηση της διατροφικής κατάστασης επειδή τα επίπεδα της λευκωματίνης στο πλάσμα εξαρτώνται από την πρόσληψη πρωτεΐνης.
Alkaline Phosphate	Οι μετρήσεις αλκαλικής φωσφατάσης χρησιμοποιούνται στη διάγνωση και τη θεραπεία της νόσου του ήπατος, των οστών και του παραθυρεοειδούς.
Aspartate Aminotransferase	Οι μετρήσεις AST χρησιμοποιούνται στη διάγνωση και θεραπεία ορισμένων τύπων ηπατικών και καρδιακών παθήσεων. Τα αυξημένα επίπεδα των τρανσαμινασών μπορούν να σηματοδοτήσουν έμφραγμα του μυοκαρδίου, ηπατική νόσο, μυϊκή δυστροφία ή βλάβη οργάνων.
Bicarbonate	Μαζί με τον προσδιορισμό του pH, χρησιμοποιούνται μετρήσεις διττανθρακικών στη διάγνωση και θεραπεία πολλών δυνητικά σοβαρών διαταραχών που σχετίζονται με ανισορροπία οξέος-βάσης στο αναπνευστικό και μεταβολικό σύστημα.

Στοιχείο	Περιγραφή
Blood Urea Nitrogen (BUN)	Οι μετρήσεις BUN χρησιμοποιούνται στη διάγνωση ορισμένων νεφρικών και μεταβολικών παθήσεων. Ο προσδιορισμός του αζώτου της ουρίας ορού είναι η πιο ευρέως χρησιμοποιούμενη δοκιμή για την αξιολόγηση της νεφρικής λειτουργίας. Σε συνδυασμό με τα επίπεδα κρεατινίνης στον ορό χρησιμοποιείται για τη διαφορική διάγνωση της προγεννητικής, νεφρικής και μετα-νεφρικής ουραιμίας. Τα υψηλά επίπεδα BUN σχετίζονται με μειωμένη νεφρική λειτουργία, αυξημένο καταβολισμό πρωτεϊνών, νεφρίτιδα, εντερική απόφραξη, απόφραξη ούρων, μεταλλική δηλητηρίαση, καρδιακή ανεπάρκεια, περιτονίτιδα, αφυδάτωση, κακοήθεια, πνευμονία, χειρουργικό σοκ, νόσο του Addison και ουραιμία. Τα χαμηλά επίπεδα BUN σχετίζονται με αμυλοείδωση, οξεία ηπατική νόσο, εγκυμοσύνη και νέφρωση. Οι φυσιολογικές διακυμάνσεις παρατηρούνται ανάλογα με την ηλικία και το φύλο ενός ατόμου, την ώρα της ημέρας διατροφικού intake - ιδιαίτερα σε σχέση με την πρόσληψη πρωτεΐνης.
Cholesterol	Ένα αυξημένο επίπεδο χοληστερόλης σχετίζεται με διαβήτη, νέφρωση, υποθυρεοειδισμό, απόφραξη των χοληφόρων και αυτές τις σπάνιες περιπτώσεις ιδιοπαθούς υπερχοληστερολαιμίας και υπερλιπιδαιμίας. Τα χαμηλά επίπεδα σχετίζονται με υπερθυρεοειδισμό, ηπατίτιδα και μερικές φορές σοβαρή αναιμία ή λοίμωξη.
Creatinine	Οι μετρήσεις κρεατινίνης είναι χρήσιμες στη διάγνωση και τη θεραπεία νεφρικών παθήσεων. Κρεατίνη φωσφοκινάση (CPK) Οι μετρήσεις της κρεατινικής κινάσης χρησιμοποιούνται στη διάγνωση και τη θεραπεία του εμφράγματος του μυοκαρδίου, των ασθενειών των σκελετικών μυών και των ασθενειών του κεντρικού νευρικού συστήματος.
Gammaglutamyl Transaminase (GGT)	Οι μετρήσεις GGT χρησιμοποιούνται κυρίως για τη διάγνωση και την παρακολούθηση της ηπατοβολικής νόσου. Αυτή τη στιγμή είναι ο πιο ευαίσθητος ενζυματικός δείκτης ηπατικής νόσου, με φυσιολογικές τιμές που σπάνια απαντώνται παρουσία ηπατικής νόσου. Χρησιμοποιείται επίσης ως ευαίσθητο τεστ διαλογής για απόκρυφο αλκοολισμό. Αυξημένα επίπεδα βρίσκονται σε ασθενείς που παίρνουν χρόνια φάρμακα όπως η φαινοβαρβιτάλη και η φαινοτοΐνη.
Globulin	Οι σφαιρίνες είναι μια διαφορετική ομάδα πρωτεϊνών που μεταφέρουν διάφορες ουσίες στο αίμα. Συμμετέχουν επίσης σε διάφορους αμυντικούς μηχανισμούς μέσα στο σώμα. Υπολογίζονται οι μετρήσεις της σφαιρίνης (Ολική πρωτεΐνη - Λευκωματίνη) και χρησιμοποιούνται για τον προσδιορισμό της συγκέντρωσης της σφαιρίνης στον ορό.
Glucose	Οι μετρήσεις γλυκόζης χρησιμοποιούνται στη διάγνωση και τη θεραπεία του καρκινώματος των παγκρεατικών νησιδίων και των διαταραχών του μεταβολισμού των υδατανθράκων, συμπεριλαμβανομένου του σακχαρώδους διαβήτη, της υπογλυκαιμίας νεογνών και της ιδιοπαθούς υπογλυκαιμίας.

Στοιχείο	Περιγραφή
Iron	Οι μετρήσεις σιδήρου (μη αίματος) χρησιμοποιούνται στη διάγνωση και θεραπεία ασθενειών, όπως αναιμία ανεπάρκειας σιδήρου, χρόνια νεφρική νόσο και αιμοχρωμάτωση (μια ασθένεια που σχετίζεται με εκτεταμένη εναπόθεση στους ιστούς δύο χρωστικών που περιέχουν σίδηρο, αιμοσιδρίνης και αιμοφουσκίνης, και χαρακτηρίζεται από μελάγχρωση του δέρματος).
Lactate Dehydrogenase (LDH)	Οι μετρήσεις LDH χρησιμοποιούνται στη διάγνωση και τη θεραπεία ηπατικών παθήσεων όπως οξεία ιική ηπατίτιδα, κίρρωση και μεταστατικό καρκίνωμα του ήπατος. καρδιακές παθήσεις, όπως έμφραγμα του μυοκαρδίου. και όγκους των πνευμόνων ή των νεφρών.
Osmolality	Οσμωτικότητα ορού είναι ένα μέτρο του αριθμού των διαλυμένων σωματιδίων σε ένα διάλυμα. Το Osmolality είναι μια υπολογισμένη τιμή στον αναλυτή χημείας: $[(1,86 * Na) + (GLUC / 18) + (BUN / 2,8) + 9]$.
Phosphorus	Υπάρχει αμοιβαία σχέση μεταξύ ασβεστίου ορού και ανόργανου φωσφόρου. Οποιαδήποτε αύξηση του επιπέδου ανόργανου φωσφόρου προκαλεί μείωση του επιπέδου ασβεστίου από έναν μηχανισμό που δεν είναι σαφώς κατανοητός. Η υπερφωσφαταιμία σχετίζεται με υπερβιταμίνωση βιταμίνης D, υποπαραθυρεοειδισμό και νεφρική ανεπάρκεια. Η υποφωσφαταιμία σχετίζεται με ραχίτιδα, υπερπαραθυρεοειδισμό και σύνδρομο Fanconi.
Potassium, Chloride, and Sodium	Οι χαμηλές τιμές χλωριούχου ορού σχετίζονται με νεφρίτιδα που χάνει το αλάτι. Addisonian κρίση, παρατεταμένος έμετος και μεταβολική οξέωση που προκαλείται από υπερβολική παραγωγή ή μειωμένη απέκκριση οξέων. Οι υψηλές τιμές χλωριούχου ορού σχετίζονται με αφυδάτωση και καταστάσεις που προκαλούν μειωμένη νεφρική ροή αίματος, όπως συμφορητική καρδιακή ανεπάρκεια. Οι μετρήσεις νατρίου χρησιμοποιούνται στη διάγνωση και θεραπεία ασθενειών που περιλαμβάνουν ανισορροπία ηλεκτρολυτών.
Total Bilirubin	Τα χαμηλά επίπεδα χολερυθρίνης σχετίζονται με απλαστική αναιμία και ορισμένους τύπους δευτερογενούς αναιμίας που προκύπτουν από τοξική θεραπεία για καρκίνωμα και χρόνια νεφρίτιδα.
Total Calcium	Οι μετρήσεις ασβεστίου χρησιμοποιούνται στη διάγνωση και τη θεραπεία της παραθυρεοειδούς νόσου, των οστών, της χρόνιας νεφρικής νόσου και της τετάνης. Η μέτρηση ασβεστίου στα ούρα χρησιμοποιείται στη διαφορική διάγνωση της υπερασβεστουρίας.
Total Protein	Οι συνολικές μετρήσεις πρωτεΐνης χρησιμοποιούνται στη διάγνωση και θεραπεία μιας ποικιλίας ασθενειών που περιλαμβάνουν το συκώτι, τα νεφρά ή τον μυελό των οστών, καθώς και άλλες μεταβολικές ή διατροφικές διαταραχές.

Στοιχείο	Περιγραφή
Triglycerides	Οι μετρήσεις τριγλυκεριδίων χρησιμοποιούνται στη διάγνωση σακχαρώδους διαβήτη, νέφρωσης, απόφραξης του ήπατος και άλλων ασθενειών που περιλαμβάνουν μεταβολισμό λιπιδίων και διάφορες ενδοκρινικές διαταραχές και στη θεραπεία ασθενών με αυτές τις ασθένειες.
Uric Acid	Οι μετρήσεις ουρικού οξέος χρησιμοποιούνται στη διάγνωση και θεραπεία πολλών νεφρικών και μεταβολικών διαταραχών, όπως νεφρική ανεπάρκεια, ουρική αρθρίτιδα, λευχαιμία, ψωρίαση, λιμοκτονία ή άλλες καταστάσεις σπατάλης και στη θεραπεία ασθενών που λαμβάνουν κυτταροτοξικά φάρμακα.

Εικόνα 2.η Βιοχημικό προφίλ

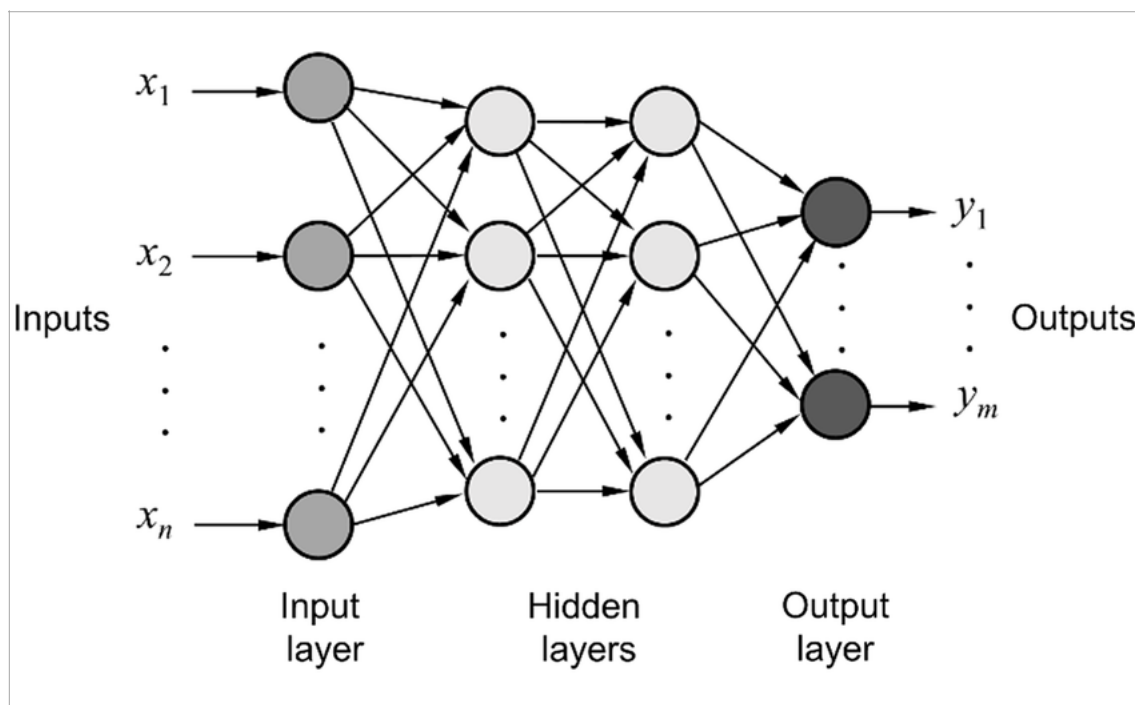
3. Νευρωνικά δίκτυα, αναγνώριση προτύπων και datasets

3.1 Multi Layer Neural Network

Τα δίκτυα πολλαπλών επιπέδων επιλύουν το πρόβλημα ταξινόμησης για μη γραμμικά σύνολα χρησιμοποιώντας κρυφά στρώματα, των οποίων οι νευρώνες δεν συνδέονται άμεσα με την έξοδο. Τα επιπρόσθετα κρυμμένα στρώματα μπορούν να ερμηνευθούν γεωμετρικά ως επιπρόσθετα υπερ-επίπεδα, τα οποία ενισχύουν την χωρητικότητα χωρισμού του δικτύου. Συνδέοντας τους τεχνητούς νευρώνες σε αυτό το δίκτυο μέσω λειτουργιών μη γραμμικής ενεργοποίησης, μπορούμε να δημιουργήσουμε πολύπλοκα, μη γραμμικά όρια αποφάσεων που μας επιτρέπουν να αντιμετωπίσουμε προβλήματα όπου οι διαφορετικές κατηγορίες δεν είναι γραμμικά διαχωρισμένες.

Η εκπαίδευση πραγματοποιείται σε εποπτευόμενο στυλ. Η βασική ιδέα είναι η παρουσίαση του διανύσματος εισόδου στο δίκτυο, υπολογίζοντας προς τα εμπρός την έξοδο κάθε επιπέδου και την τελική έξοδο του δικτύου. Για το επίπεδο εξόδου είναι γνωστές οι επιθυμητές τιμές και επομένως τα βάρη μπορούν να ρυθμιστούν όπως για ένα δίκτυο μονής στρώσης στην περίπτωση του αλγόριθμου BP σύμφωνα με τον κανόνα της διαβάθμισης.

Για τον υπολογισμό των αλλαγών βάρους στο κρυφό στρώμα, το σφάλμα στο επίπεδο εξόδου μεταδίδεται ξανά σε αυτά τα επίπεδα ανάλογα με τα βάρη σύνδεσης. Αυτή η διαδικασία επαναλαμβάνεται για κάθε δείγμα στο σετ εκπαίδευσης. Ένας κύκλος μέσω του σετ κατάρτισης ονομάζεται εποχή. Ο αριθμός των εποχών που απαιτούνται για την εκπαίδευση του δικτύου εξαρτάται από διάφορες παραμέτρους, ειδικά από το σφάλμα που υπολογίζεται στο επίπεδο εξόδου.



Εικόνα 3.α Πολυεπίπεδο νευρωνικό δίκτυο

3.2 Περιβάλλον υλοποίησης

Scikit-learn python

Το Scikit-learn (πρώην scikits.learn και επίσης γνωστό ως sklearn) είναι μια βιβλιοθήκη μηχανικής εκμάθησης δωρεάν λογισμικού για τη γλώσσα προγραμματισμού Python. [3] Διαθέτει διάφορους αλγόριθμους ταξινόμησης, παλινδρόμησης και ομαδοποίησης, συμπεριλαμβανομένων μηχανών φορέα υποστήριξης, τυχαίων δασών, ενίσχυσης κλίσης, k-μέσων και DBSCAN, και έχει σχεδιαστεί για να λειτουργεί με τις αριθμητικές και επιστημονικές βιβλιοθήκες Python NumPy και SciPy.

Γλωσσάρι προγραμματιστικών εννοιών και βιβλιοθήκες

- **StandardScaler**

κάνει το μέσο όρο της κατανομής 0. Περίπου το 68% των τιμών θα είναι μεταξύ -1 και 1.

- **sklearn.model_selection.train_test_split**

Διαχωρισμός σε πίνακες σε τυχαία υποσύνολα εκπαίδευσης και δοκιμής

παράδειγμα χρήσης

```
X_train, X_test, y_train, y_test = train_test_split(predict_dataset, target_dataset,
test_size=0.1, random_state=42)
```

```
# Trains scaler which standarizes all the features in a binary mode
```

```
sc = StandardScaler()
```

```
sc.fit(X_train)
```

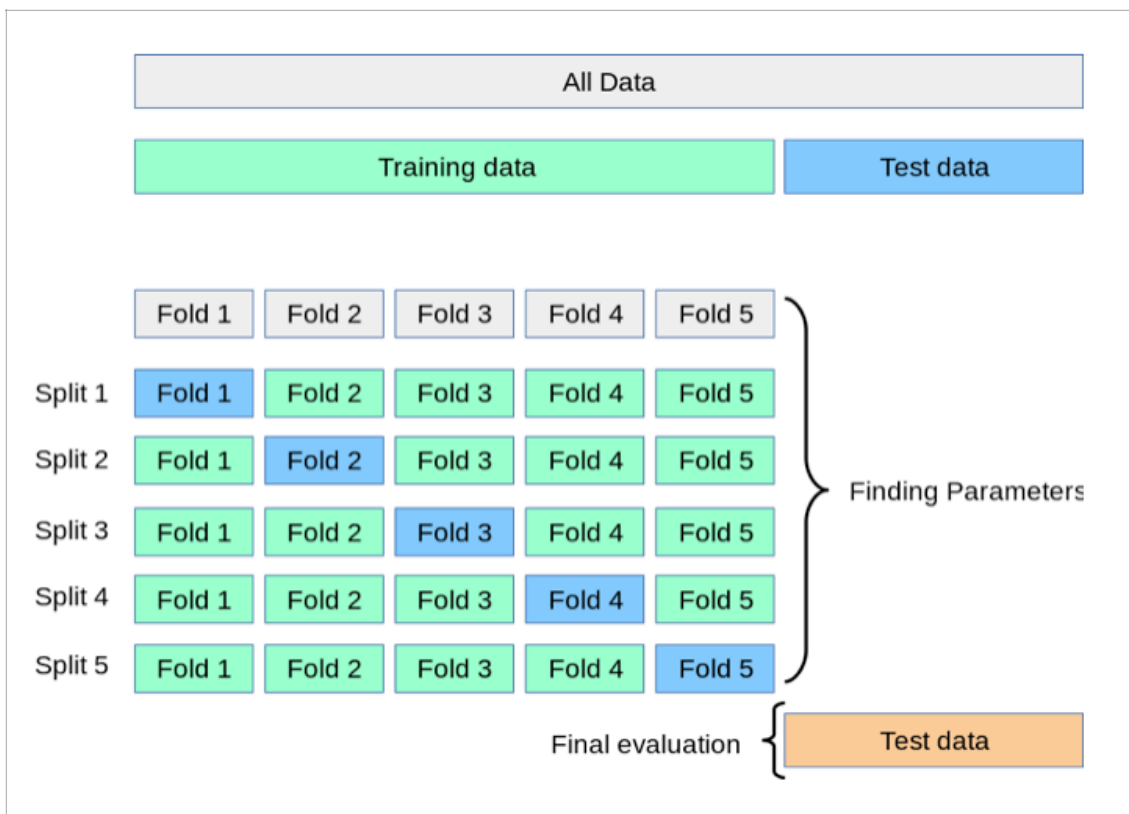
```
# Apply the scaler to the X training data
```

```
X_train =sc.transform(X_train)
```

- **Cross validation**

.Σε ένα πρόβλημα πρόβλεψης, ένα μοντέλο λαμβάνει συνήθως ένα σύνολο δεδομένων γνωστών δεδομένων στο οποίο εκτελείται η εκπαίδευση (σύνολο δεδομένων εκπαίδευσης) και ένα σύνολο δεδομένων άγνωστων δεδομένων (ή δεδομένα που έχουν δει πρώτη φορά) έναντι των οποίων δοκιμάζεται το μοντέλο (ονομάζεται σύνολο δεδομένων επικύρωσης ή δοκιμή σετ). Ο στόχος της διασταυρούμενης επικύρωσης είναι να δοκιμάσει την ικανότητα του μοντέλου να προβλέψει νέα δεδομένα που δεν χρησιμοποιήθηκαν για την εκτίμησή του, προκειμένου να επισημανθούν προβλήματα όπως η υπερβολική τοποθέτηση ή η μεροληψία επιλογής και να δοθεί μια εικόνα για το πώς το μοντέλο θα γενικευτεί σε ένα ανεξάρτητο σύνολο δεδομένων (δηλαδή, ένα άγνωστο σύνολο δεδομένων, για παράδειγμα από ένα πραγματικό πρόβλημα).

Ένας κύκλος διασταυρούμενης επικύρωσης περιλαμβάνει τον διαχωρισμό ενός δείγματος δεδομένων σε συμπληρωματικά υποσύνολα, την εκτέλεση της ανάλυσης σε ένα υποσύνολο (που ονομάζεται εκπαιδευτικό σύνολο) και την επικύρωση της ανάλυσης στο άλλο υποσύνολο (που ονομάζεται σύνολο επικύρωσης ή σύνολο δοκιμών). Για να μειωθεί η μεταβλητότητα, στις περισσότερες μεθόδους εκτελούνται πολλαπλοί γύροι εγκυρότητας με χρήση διαφορετικών κατατμήσεων και τα αποτελέσματα επικύρωσης συνδυάζονται (π.χ. κατά μέσο όρο) στους γύρους για να δώσουν μια εκτίμηση της προγνωστικής απόδοσης του μοντέλου.



Εικόνα 3.β παράδειγμα χρήσης cross validation με πενταπλό διαχωρισμό των δεδομένων και αξιολόγηση του μοντέλου.

```
TestScores = cross_val_score(ppn, X_test, np.ravel(y_test),cv=5)
```

• **random state**

Ελέγχει την αναδιάταξη που εφαρμόζεται στα δεδομένα πριν από την εφαρμογή του διαχωρισμού.

• **solver**

Αλγόριθμος για χρήση στο πρόβλημα βελτιστοποίησης.

• **hidden layer size tuple, length = n_layers - 2, default=(100,)**

Το n στοιχείο αντιπροσωπεύει τον αριθμό των νευρώνων στο κρυφό στρώμα.

• **activation**

Λειτουργία ενεργοποίησης για τα hidden layers .

«Ταυτότητα», ενεργοποίηση χωρίς λειτουργία, χρήσιμη για την υλοποίηση γραμμικής συμφόρησης, επιστροφές $f(x) = x$

«Logistic», η συνάρτηση logistic sigmoid, επιστρέφει $f(x) = 1 / (1 + \exp(-x))$.

«Tanh», η συνάρτηση υπερβολικού μαύρου, επιστρέφει $f(x) = \tanh(x)$.

«Relu», η λειτουργία διορθωμένης γραμμικής μονάδας, επιστρέφει $f(x) = \max(0, x)$

• **alpha**

Παράμετρος ποινής L2 (όρος κανονικοποίησης).

• **max_iter**

Μέγιστος αριθμός επαναλήψεων Ο επιλυτής επαναλαμβάνεται έως τη σύγκλιση

• **RadViz visualizer**

Το RadViz είναι ένας αλγόριθμος οπτικοποίησης πολυπαραγοντικών δεδομένων που απεικονίζει κάθε διάσταση χαρακτηριστικών γύρω από την περιφέρεια ενός κύκλου και στη συνέχεια σχεδιάζει σημεία στο εσωτερικό του κύκλου έτσι ώστε το σημείο να ομαλοποιεί τις τιμές του στους άξονες από το κέντρο σε κάθε τόξο.

3.2 Datasets και σχεδιασμός

Τα datasets τα οποία θα χρησιμοποιηθούν προέρχονται από το NHANES. Το National Health and Nutrition Examination Survey (NHANES) είναι ένα πρόγραμμα μελετών που αποσκοπεί στην αξιολόγηση της υγείας και της διατροφικής κατάστασης των ενηλίκων και των παιδιών στις Ηνωμένες Πολιτείες. Η έρευνα είναι μοναδική, καθώς συνδυάζει συνεντεύξεις και φυσικές εξετάσεις. Το NHANES είναι ένα σημαντικό πρόγραμμα του Εθνικού Κέντρου Στατιστικών Υγείας (NCHS). Το NCHS είναι μέρος των Κέντρων Ελέγχου και Πρόληψης Νοσημάτων (CDC) και έχει την ευθύνη για την παραγωγή ζωτικών υγειονομικών στατιστικών για την Αμερική. Το πρόγραμμα NHANES ξεκίνησε στις αρχές της δεκαετίας του 1960 και διεξήχθη ως μια σειρά ερευνών που εστιάζουν σε διαφορετικές ομάδες πληθυσμού σε θέματα υγείας. Το 1999, μεταλλάχθηκε σε ένα συνεχές πρόγραμμα που έχει μια μεταβαλλόμενη εστίαση σε μια ποικιλία μετρήσεων υγείας και διατροφής για την κάλυψη αναδυόμενων αναγκών. Το πρόγραμμα εξετάζει ένα εθνικό αντιπροσωπευτικό δείγμα περίπου 5.000 ατόμων κάθε χρόνο.

Πιο συγκεκριμένα βρέθηκαν και χρησιμοποιήθηκαν τα datasets του 2013-2014 και περιλαμβάνουν δημογραφικά στοιχεία, εξετάσεις(αίματος, φυσιολογίας κλπ) , δίαιτας, εργαστηριακές εξετάσεις, ερωτηματολόγιο και ιατρικό ιστορικό, φαρμακευτικό ιστορικό περισσότερες λεπτομέρειες υπάρχουν στο appendix και σχετικές παραπομπές στη βιβλιογραφία.

Ο σχεδιασμός και η λογική της ανάλυσης των παραπάνω δεδομένων για την παρούσα μελέτη **βασίζεται στην υπόθεση ότι στοιχεία παρατήρησης**(ιατρικό ιστορικό, διαιτητικές συνήθειες, lifestyle κλπ) ή/και στοιχεία εργαστηριακά μπορούν να **δώσουν μια εικόνα της κατάστασης ενός ατόμου**. Καθώς το **βάρος** είναι μια σημαντική ένδειξη της υγείας, θα θεωρηθεί το BMI(δείκτης ύψους/μάζας) ως στόχος πρόβλεψης(target) και τα υπόλοιπα στοιχεία θα χρησιμοποιηθούν ως χαρακτηριστικά στοιχεία για την εκπαίδευση του μοντέλου πρόβλεψης(features/characteristics).

Βάση του BMI θα γίνει ένα binary classification των ομάδων σε λιποβαρής, κανονικούς και υπέρβαρους. Η τελευταία κατηγορία λόγω σχετικά περιορισμένου δείγματος θα περιλαμβάνει και τους παχύσαρκους. Κατόπιν τα αποτελέσματα θα αξιολογούνται και θα “καλιμπράρονται” για τη βελτιστοποίηση της ευστοχίας του νευρωνικού δικτύου.

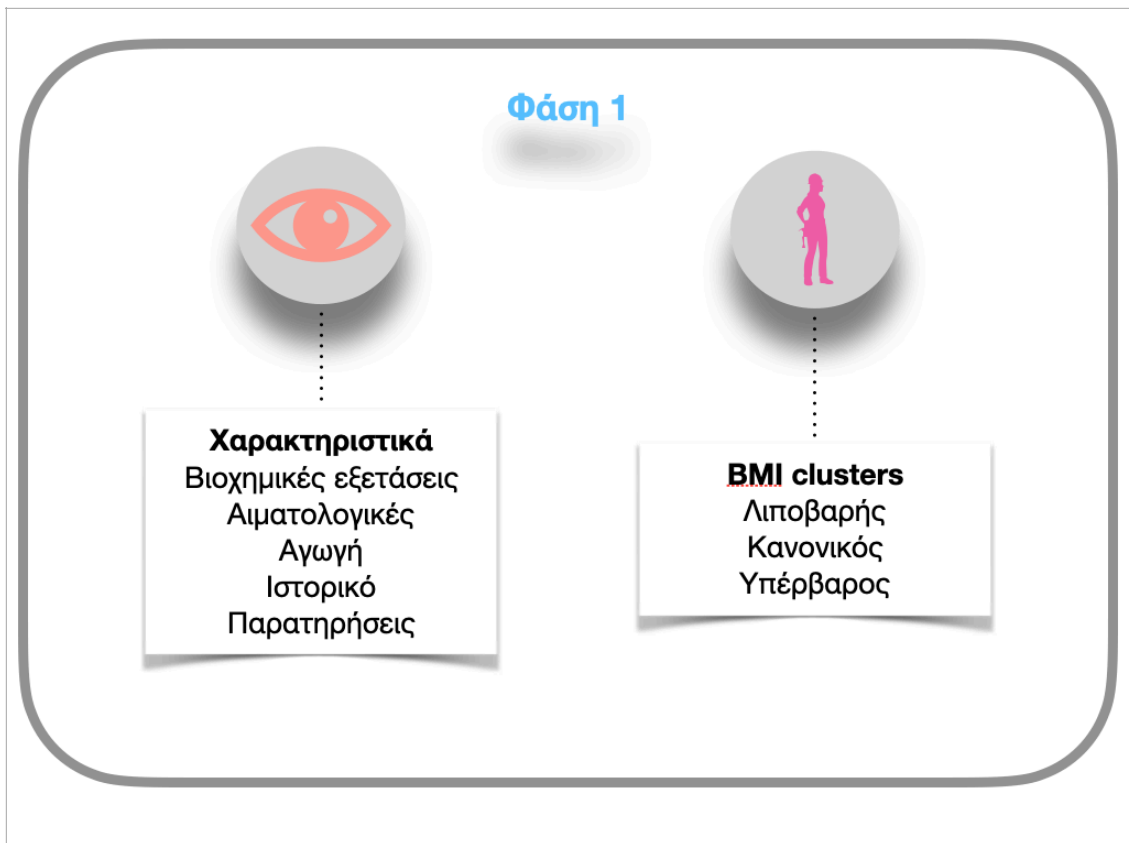
Κατόπιν θα εντοπιστούν στοιχεία από τα χαρακτηριστικά τα οποία αν είναι πιο έντονα σε άτομα με κανονικό βάρος θα χρησιμοποιηθούν με μια ανάλογη διαδικασία ως στόχοι (targets) πρόβλεψης ενός νευρωνικού δικτύου με χαρακτηριστικά(features/inputs) το διαιτητικό προφίλ του δείγματος.

Αν το νευρωνικό δίκτυο έχει αποτελεσματικότητα πρόβλεψης τότε αποδεικνύεται η υπόθεση ότι το βάρος προϋπολογίζεται από τα αρχικά χαρακτηριστικά, εξετάσεις κλπ και τα χαρακτηριστικά αυτά συνδέονται με τις διατροφικές συνήθειες και τις διατροφικές εισροές.

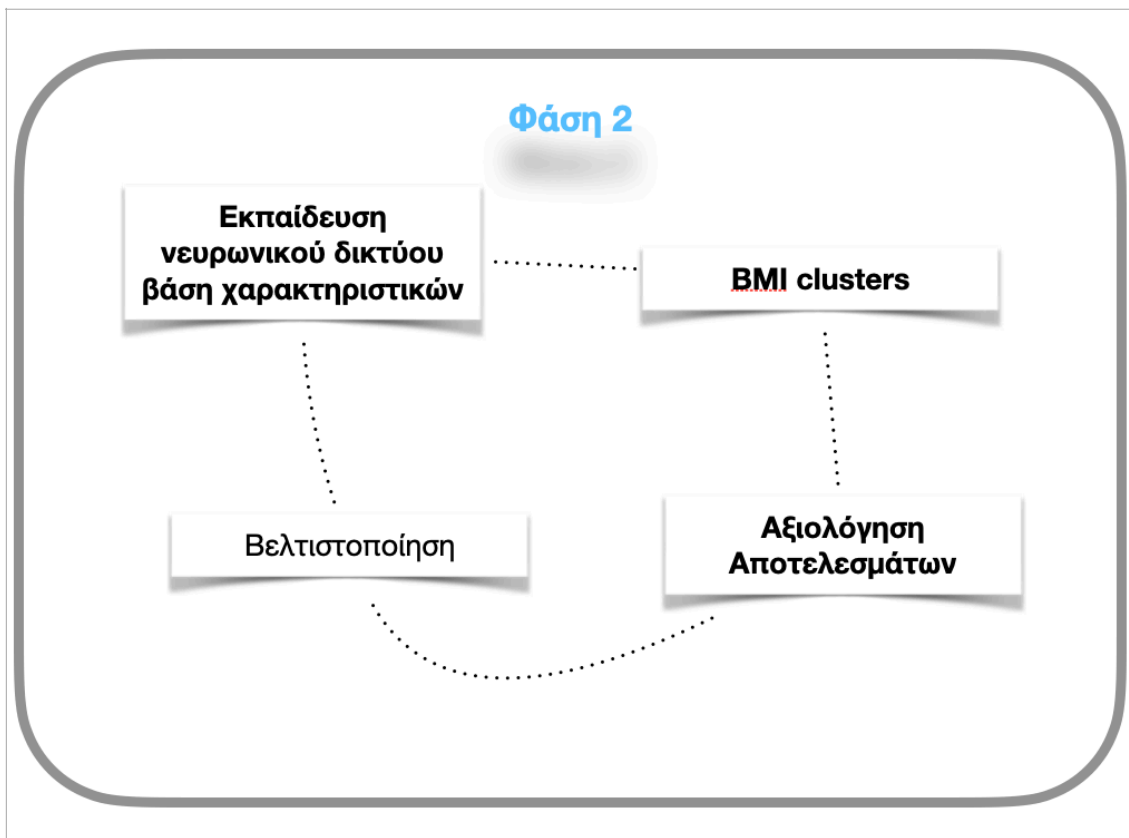
Συγκεκριμένα :

if	bmi	less	than	18.5	underweight
else if	bmi	between	18.5	25	normal
else if	bmi	over	25		overweight

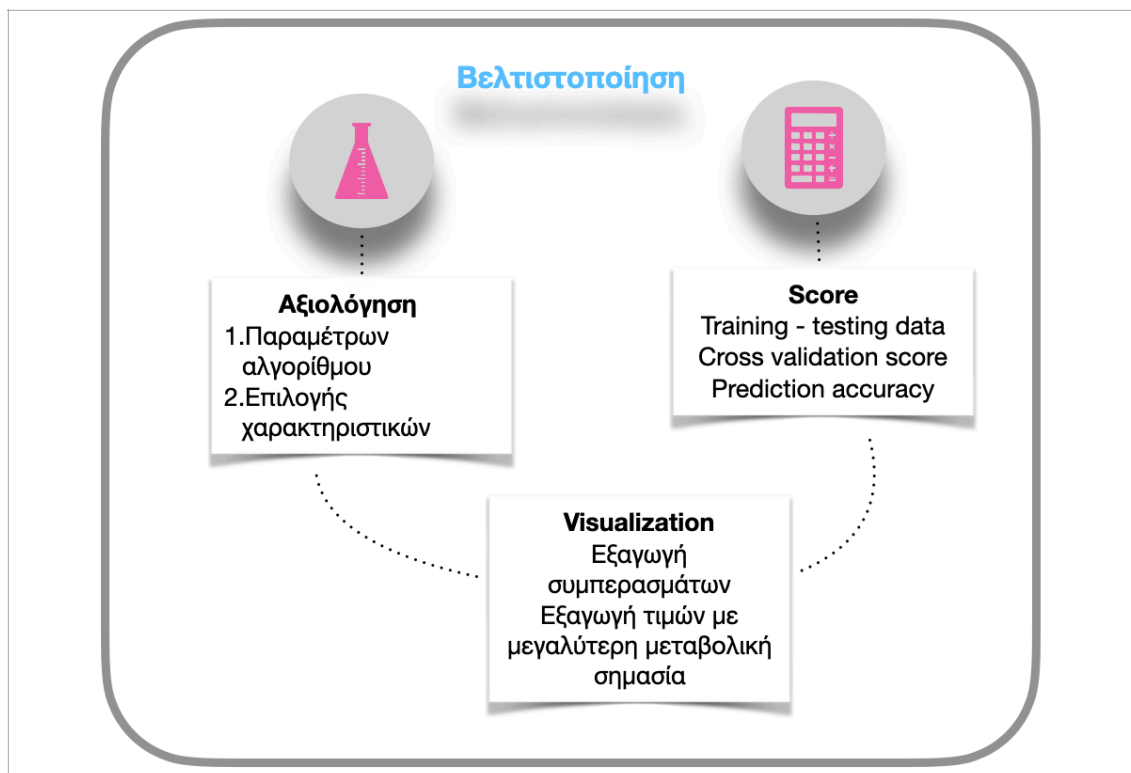
Παρακάτω ακολουθεί μια εννοιολογική ανάλυση της διαδικασίας σκέψης και υλοποίησης του προγράμματος.



Εικόνα 3.γ Conceptual analysis 1



Εικόνα 3.δ Conceptual analysis 2



Εικόνα 3.ε Conceptual analysis 3

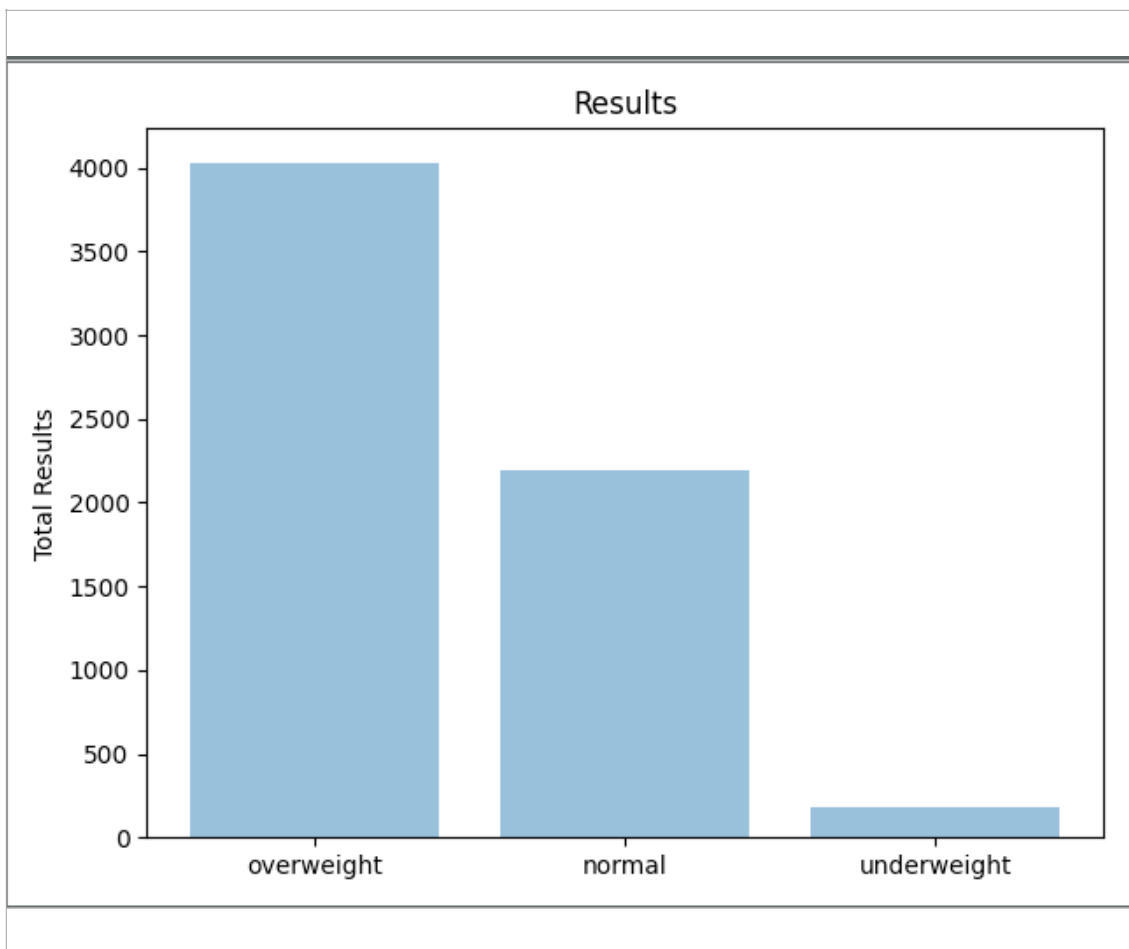
4. Υλοποίηση - Αποτελέσματα

4.1 Ανάλυση δεδομένων Α'

Θα χρησιμοποιηθούν τα διαθέσιμα δεδομένα και θα διαχωριστούν σε 3 κατηγορίες σύμφωνα με το BMI. Οι κατηγορίες θα είναι λιπόβαρής, κανονικό βάρος, υπέρβαροι. Η διαδικασία θα πραγματοποιηθεί με ένα for loop σε περιβάλλον python.

```
for row in target:
    if (row > 25):
        r = np.append(r, [[1, 0, 0]], axis=0)
    elif (row > 18 and row <= 25):
        r = np.append(r, [[0, 1, 0]], axis=0)
    elif (row <= 18):
        r = np.append(r, [[0, 0, 1]], axis=0)
target_dataset = pd.DataFrame({'overweight': r[:, 0], 'normal': r[:, 1],
'underweight': r[:, 2]})
```

Τα αποτελέσματα καταγράφονται παρακάτω:



Εικόνα 4.α Data analysis

4.2 Νευρωνικό δίκτυο Α'

Στη συνέχεια θα σχεδιαστεί το νευρωνικό δίκτυο με στόχο την πρόβλεψη του BMI λαμβάνοντας ως input τα χαρακτηριστικά των βιοχημικών εξετάσεων. Έγιναν δοκιμές οι οποίες συμπεριέλαβαν γενικές αίματος, γενικές αίματος και κάπνισμα γενικές αίματος και χοληστερόλη, γενικές αίματος και χοληστερόλη και γλυκοαιμοσφαιρίνη αλλά το αποτέλεσμα δεν βελτιώθηκε για αυτόν τον λόγο προεκρίθη ως διαδικασία η χρήση μόνο των βιοχημικών εξετάσεων, καθώς δίνει καλύτερα αποτελέσματα με μικρότερο κόστος.

Data for model training	f0	f1	f2	f3	f4	f5	... f37
0	129.0	16.0	16.0	10.0	3.57	27.0	...

Data for model training	f0	f1	f2	f3	f4	f5	... f37
1	97.0	18.0	29.0	16.0	5.71	23.0	...
2	99.0	22.0	16.0	14.0	5.00	23.0	...
3	78.0	36.0	28.0	31.0	11.07	31.0	...
4	95.0	24.0	16.0	18.0	6.43	25.0	...
...
6408	46.0	24.0	47.0	12.0	4.28	29.0	...
6409	64.0	24.0	23.0	17.0	6.07	28.0	...
6410	43.0	20.0	12.0	26.0	9.28	25.0	...
6411	75.0	27.0	26.0	13.0	4.64	25.0	...
6412	65.0	26.0	13.0	10.0	3.57	23.0	...
[6413 rows x 35 columns]							

Πίνακας 4.α Data inputs

ytrain	overweight	normal	underweight
5211	1.0	0.0	0.0
2188	0.0	0.0	1.0
2077	1.0	0.0	0.0
5341	0.0	1.0	0.0
2120	0.0	1.0	0.0
...
5032	1.0	0.0	0.0
2106	1.0	0.0	0.0
452	1.0	0.0	0.0
3438	1.0	0.0	0.0

ytrain	overweight	normal	underweight
5886	1.0	0.0	0.0
[6413 rows x 35 columns]			

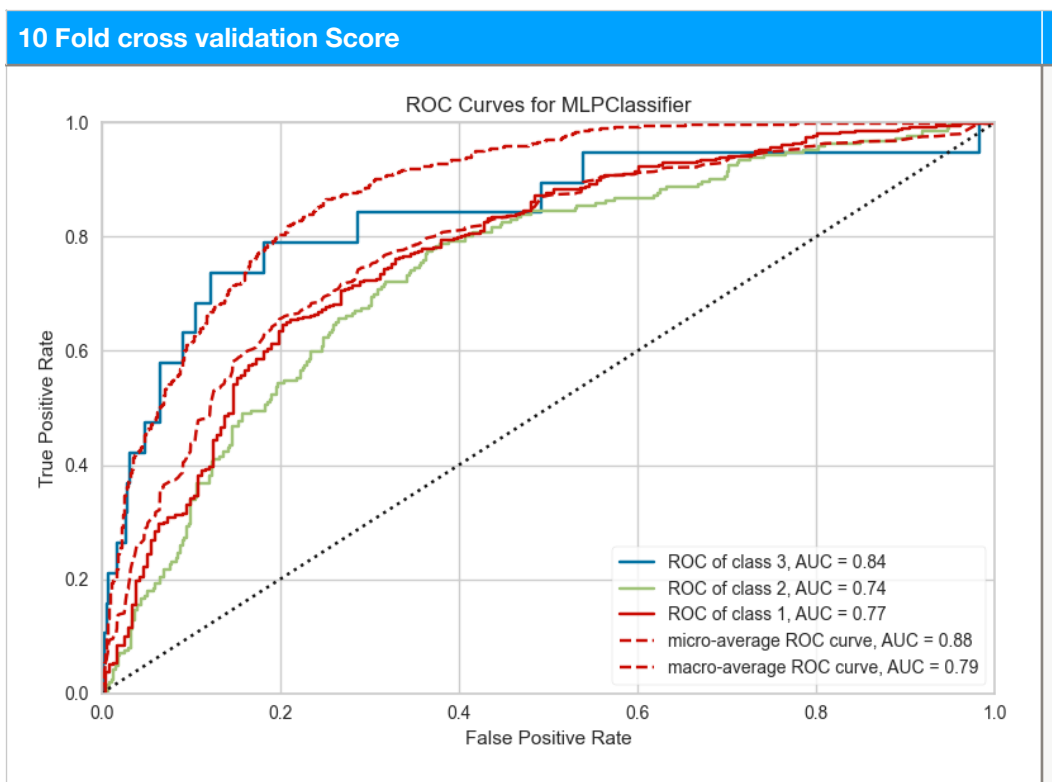
Πίνακας 4.β Data inputs

Testing	score:	0.75545171339563 87
Training	score:	0.73505458326113 33
accuracy	score	0.75545171339563 87

Πίνακας 4.γ mlp.scores & accuracy score

10 Fold cross validation Score

Training
[0.7266436 0.70363951 0.7677643 0.72963605 0.71577123 0.62218371 0.59618718 0.59445407 0.73310225 0.75563258]
Training average
0.6945014482498064
Testing
[0.69230769 0.70769231 0.75 0.625 0.703125 0.578125 0.703125 0.671875 0.65625 0.5625]
Testing Average
0.665

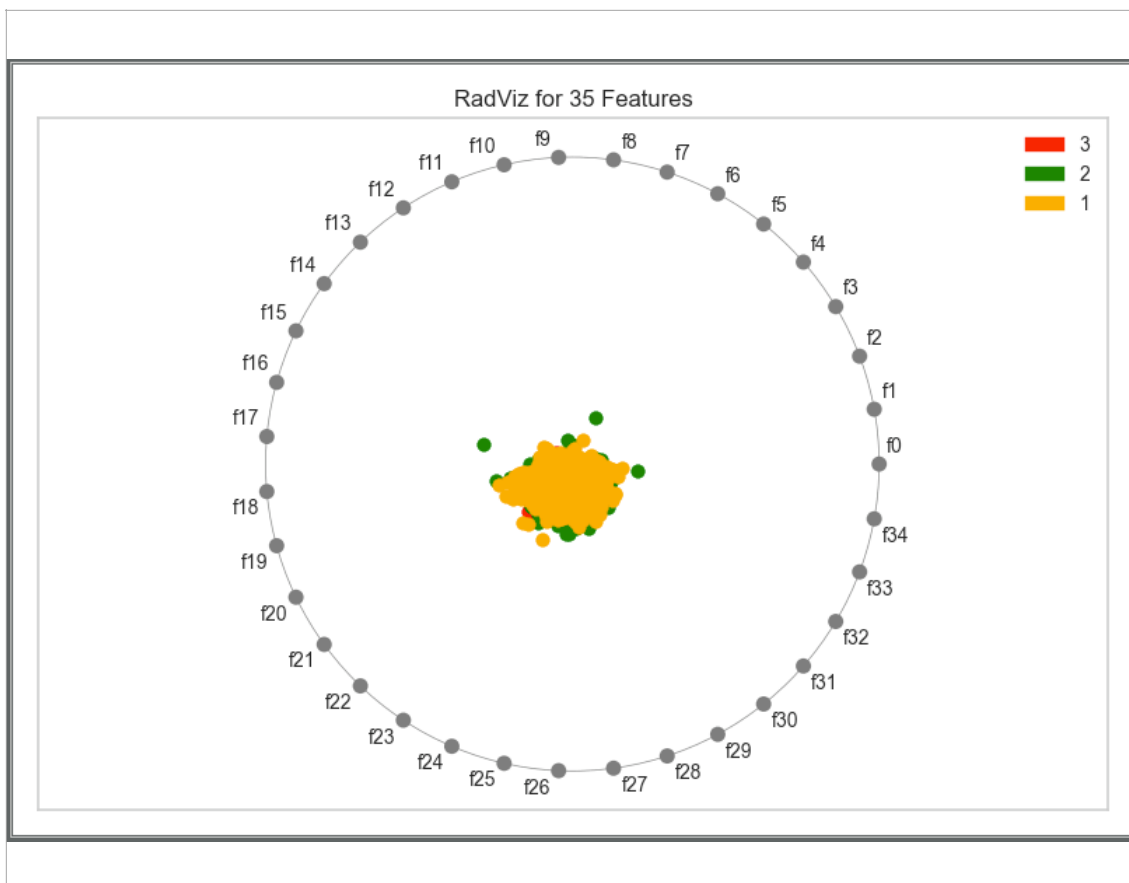


Πίνακας 4.δ 10-Fold cross validation Score

Το νευρωνικό δίκτυο έχει ένα testing average σε 10 fold cross validation 0.665 το οποίο είναι σημαντικό σε ανομοιογενές δείγμα μόλις 6500 περίπου ατόμων. Παρακάτω αναγράφεται ο αλγόριθμος με τα χαρακτηριστικά που έχει δώσει τα παραπάνω αποτελέσματα.

```
mlp = LabelPowerSet(
    MLPClassifier(hidden_layer_sizes=(4,), activation='logistic', solver='adam',
alpha=1e-5, learning_rate='adaptive', max_iter=1500))
```

Στη συνέχεια χρησιμοποιείται ο RadViz αλγόριθμος οπτικοποίησης πολυπαραγοντικών δεδομένων. Τα αποτελέσματα παρατίθενται παρακάτω. Χωρίζονται τα χαρακτηριστικά σε 3 κατηγορίες αναφορικά με το βάρος και τοποθετούνται στον κύκλο. Η κατηγορία 2 είναι το κανονικό βάρος.



Πίνακας 4.ε Rad viz /35 features

Όπως φαίνεται παραπάνω στο κανονικό βάρος ξεχωρίζουν 3 στοιχεία τα οποία είναι παρακάτω

LBXSC3SI	Bicarbon ate (mmol/L)	Μορφή διοξειδίου του άνθρακα, στο αίμα Το διττανθρακικό άλας, επίσης γνωστό ως HCO ₃ , είναι ένα υποπροϊόν του μεταβολισμού του σώματός. Το αίμα φέρνει όξινο ανθρακικό στους πνεύμονές και στη συνέχεια εκπνέεται ως διοξείδιο του άνθρακα. Τα νεφρά βοηθούν επίσης στη ρύθμιση των διττανθρακικών. Το διττανθρακικό απεκκρίνεται και απορροφάται από τα νεφρά. Αυτό ρυθμίζει το pH του σώματος ή την ισορροπία οξέος.
----------	-----------------------	---

LBXSASSI	Aspartate aminotransferase AST (IU/L)	Η ασπαρτική αμινοτρανσφεράση (AST) είναι ένα ένζυμο που βρίσκεται σε κύτταρα σε όλο το σώμα αλλά κυρίως στην καρδιά και στο ήπαρ και, σε μικρότερο βαθμό, στους νεφρούς και τους μύες. Σε υγιή άτομα, τα επίπεδα AST στο αίμα είναι χαμηλά. Όταν τραυματίζονται τα κύτταρα του ήπατος ή των μυών, απελευθερώνουν AST στο αίμα.
LBXSGI	Glucose, refrigerated serum (mg/dL)	Η γλυκόζη είναι μια απλή ζάχαρη με τον μοριακό τύπο C ₆ H ₁₂ O ₆ . Η γλυκόζη είναι ο πιο άφθονος μονοσακχαρίτης, μια υποκατηγορία υδατανθράκων.

Πίνακας 4.ζ Βιοχημικά

**Σημείωση: Εξετάστηκαν και τα κοντινά στοιχεία στον κύκλο αλλά αποκλείστηκαν λόγω του ότι σύμφωνα με την βιβλιογραφία είναι μάλλον απίθανο να συνδέονται άμεσα με τον μεταβολισμό ή το βάρος.

Στη συνέχεια θα χρησιμοποιηθούν τα παραπάνω στοιχεία και θα κατηγοριοποιηθούν ανάλογα με το εύρος τιμών, δηλαδή εντός κανονικότητας και εκτός κανονικότητας (πάνω ή κάτω από το κανονικό εύρος βλέπε παράρτημα). Θα χρησιμοποιηθούν ως στόχος και θα εκπαιδευτεί ένα νευρωνικό δίκτυο το οποίο θα προβλέπει τις τιμές των παραπάνω λαμβάνοντας ως εισροές το διαιτητικό προφίλ.

4.1 Ανάλυση δεδομένων Β'

Θα χρησιμοποιηθούν τα βιοστοιχεία τα οποία εντοπίστηκαν από το προηγούμενο δίκτυο και θα διαχωριστούν σε 2 κατηγορίες σύμφωνα με το εύρος τιμών. Οι κατηγορίες θα είναι κανονικές και μη κανονικές. Η διαδικασία θα πραγματοποιηθεί με σε περιβάλλον python με ένα sql query.

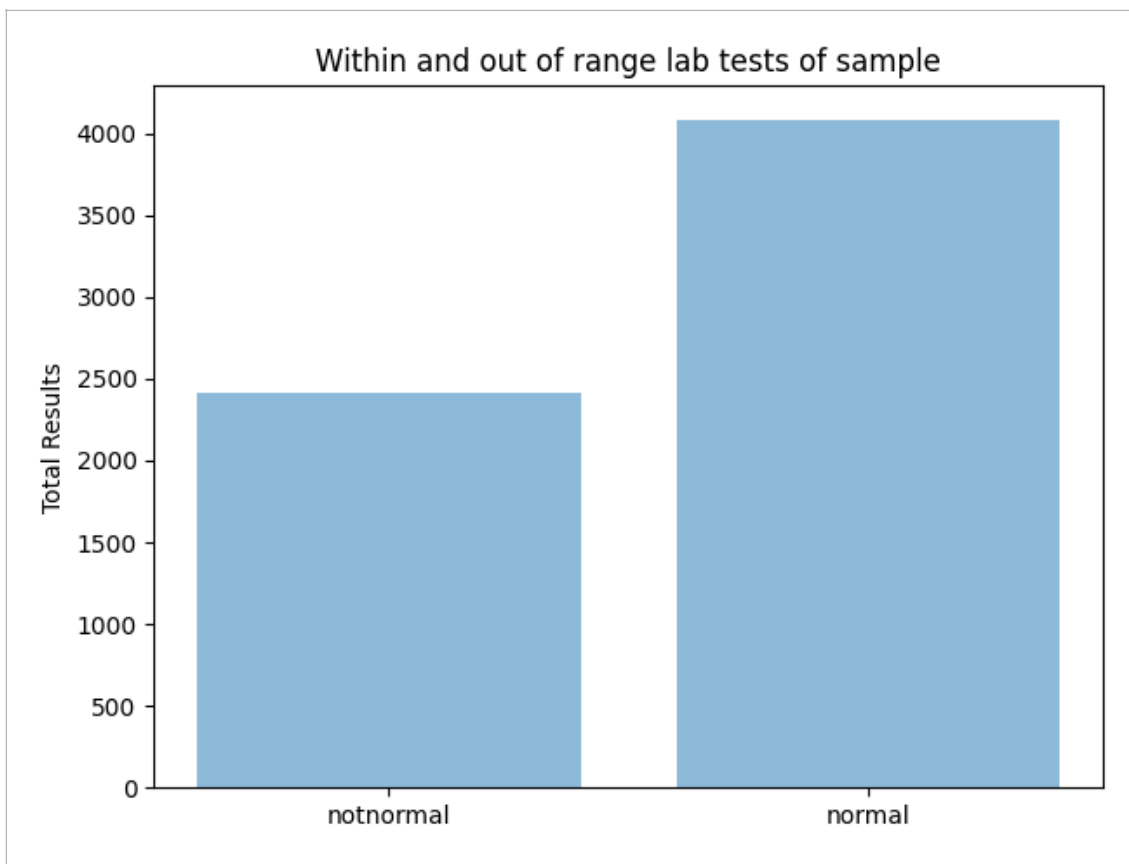
```
select LBXSASSI,LBXSC3SI,LBXSGI from labs join examination on
examination.SEQN=labs.SEQN
where LBXSASSI>=36 AND LBDSCASI not null and LBDSGBSI not null and
LBXSASSI not null
UNION select LBXSAPSI,LBXSC3SI,LBXSGI from labs join examination on
examination.SEQN=labs.SEQN
where LBXSASSI<=10 AND LBDSCASI not null and LBDSGBSI not null and
LBXSASSI not null
UNION select LBXSAPSI,LBXSC3SI,LBXSGI from labs join examination on
examination.SEQN=labs.SEQN
```



```

where LBXSC3SI<=22 AND LBDSCASI not null and LBDSGBSI not null and
LBXSASSI not null
union select LBXSC3SI,LBXSC3SI,LBXSGL from labs join examination on
examination.SEQN=labs.SEQN
where LBXSC3SI>=29 AND LBDSCASI not null and LBDSGBSI not null and
LBXSASSI not null
union select LBXSASSI,LBXSC3SI,LBXSGL from labs join examination on
examination.SEQN=labs.SEQN
where LBXSGL<70 AND LBDSCASI not null and LBDSGBSI not null and
LBXSASSI not null
union select LBXSASSI,LBXSC3SI,LBXSGL from labs join examination on
examination.SEQN=labs.SEQN
where LBXSGL>110 AND LBDSCASI not null and LBDSGBSI not null and
LBXSASSI not null
    
```

Τα αποτελέσματα παρατίθενται παρακάτω:



Πίνακας 4.η Data analysis β'

Κατόπιν θα διαμορφωθεί ένα νευρωνικό δίκτυο το οποίο λαμβάνει ως input μια σειρά παραγόντων δίαιτας τα οποία αναγράφονται στο παράρτημα .

4.2 Νευρωνικό δίκτυο Β'

Στη συνέχεια θα σχεδιαστεί το νευρωνικό δίκτυο με στόχο την πρόβλεψη των παραπάνω στοιχείων δηλαδή των τιμών των Bicarbonate (mmol/L), Aspartate aminotransferase AST (IU/L), Glucose, refrigerated serum (mg/dL) λαμβάνοντας ως input τα χαρακτηριστικά των στοιχείων δίαιτας όπως αναγράφεται στο παράρτημα.

Data for model training	f0	f1	f2	f3	f4 f72
0	4.0		2.0	2.0	11.0	1574.0	...
1	3.0		1.0	2.0	8.0	5062.0	...
2	2.0		2.0	1.0	27.0	1743.0	...
3	4.0		2.0	1.0	7.0	1421.0	...
4	3.0		1.0	2.0	17.0	1785.0	...
...
5948	2.0		2.0	2.0	15.0	1225.0	...
5949	4.0		2.0	2.0	14.0	2181.0	...
5950	1.0		2.0	2.0	24.0	2432.0	...
5951	3.0		2.0	2.0	24.0	4687.0	...
5952	3.0		2.0	1.0	31.0	2475.0	...
[5953 rows x 73 columns]							

Πίνακας 4.2.α Data inputs

ytrain	outofRange		normal
612	0.0		1.0
4456	0.0		1.0
4949	0.0		1.0
3082	1.0		0.0
4759	1.0		0.0
...		...	
2272	1.0		0.0
2047	0.0		1.0
2306	1.0		0.0
5267	0.0		1.0
1409	1.0		0.0
[5357 rows x 2 columns]			

Πίνακας 4.2.β Data training

Testing	score:	0.6157718120805369
Training	score:	0.632070188538361
accuracy	score	0.6157718120805369

Πίνακας 4.2.β mlp.scores & accuracy score

10 Fold cross validation Score	

Training	
[0.6511194 0.60447761 0.60447761 0.63432836 0.62873134 0.61380597 0.63059701 0.67850467 0.63364486 0.63925234]	
Training average	
0.6318939182591714	

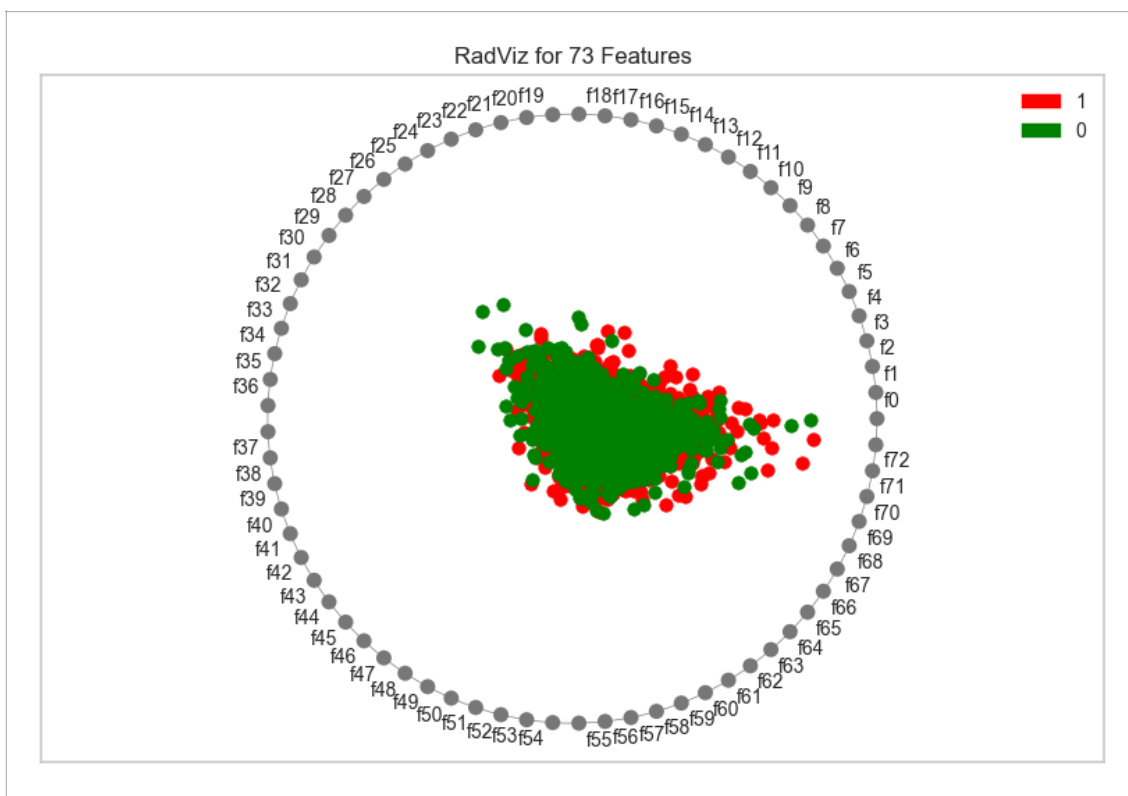
10 Fold cross validation Score	
Testing	
	[0.58333333 0.73333333 0.5 0.61666667 0.66666667 0.58333333 0.6440678 0.62711864 0.55932203 0.55932203]
Testing Average	
	0.607316384180791

Πίνακας 4.2.γ 10-fold cross validation

Το υλοποιημένο δίκτυο έχει ένα testing average σε 10 fold cross validation 0.607316384180791 το οποίο είναι σημαντικό σε σχετικά ανομοιογενές δείγμα μόλις 6000 περίπου ατόμων. Παρακάτω αναγράφεται ο αλγόριθμος με τα χαρακτηριστικά που έχει δώσει τα παραπάνω αποτελέσματα.

```
mlp = LabelPowerSet( MLPClassifier(hidden_layer_sizes=(6,),
activation='logistic', solver='adam', alpha=1e-5, learning_rate='adaptive',
max_iter=1500))
```

Στη συνέχεια χρησιμοποιείται ο RadViz αλγόριθμος οπτικοποίησης πολυπαραγοντικών δεδομένων. Τα αποτελέσματα παρατίθενται παρακάτω. Χωρίζονται τα χαρακτηριστικά σε 2 κατηγορίες αναφορικά με τους διατροφικούς παράγοντες και τοποθετούνται στον κύκλο. Πάλι κάποιοι παράγοντες ξεχωρίζουν τόσο στα άτομα που ανήκουν στην ομάδα 1 (δείκτες μη κανονικοί) όσο και στην ομάδα 0 (δείκτες εντός κανονικού εύρους).



Πίνακας 4.2.δ Rad viz /35 features

f26	DR1TVB6	Vitamin B6 (mg)	Η βιταμίνη B6, επίσης γνωστή ως πυριδοξίνη, είναι μια υδατοδιαλυτή βιταμίνη που χρειάζεται το σώμα σας για διάφορες λειτουργίες. Είναι σημαντικό για τον μεταβολισμό πρωτεϊνών, λιπών και υδατανθράκων και τη δημιουργία ερυθρών αιμοσφαιρίων και νευροδιαβιβαστών.
f28	DR1TFA	Folic acid (mcg)	Το φολικό οξύ είναι απαραίτητο για την ανάπτυξη των κυττάρων. για το μεταβολισμό συγκεκριμένων βιοχημικών αντιδράσεων στο σώμα, όπως η μετατροπή της ομοκυστεΐνης σε μεθειονίνη. και για το μεταβολισμό συγκεκριμένων αντισπασμωδικών φαρμάκων.
f31	DR1TCHL	Total choline (mg)	Η χολίνη είναι ένα βασικό θρεπτικό συστατικό που υπάρχει φυσικά σε ορισμένα τρόφιμα και διατίθεται ως συμπλήρωμα διατροφής. Η χολίνη είναι μια πηγή μεθυλικών ομάδων που απαιτούνται για πολλά στάδια στο μεταβολισμό. Ο οργανισμός χρειάζεται χολίνη για να συνθέσει φωσφατιδυλοχολίνη και σφιγγομυελίνη, δύο κύρια φωσφολιπίδια ζωτικής σημασίας για τις κυτταρικές μεμβράνες.

5. Επίλογος

5.1 Πρόληψη

Η ελαχιστοποίηση της εμφάνισης ασθενειών είναι η καλύτερη αντιμετώπιση τους. Σε συνδυασμό με την καλύτερη θεραπεία και φροντίδα η πρόληψη μπορεί να αποτελέσει και σε πολλές περιπτώσεις ήδη αποτελεί καθοριστικό παράγοντα.

Η πρόληψη μπορεί να πραγματοποιηθεί σε διάφορα επίπεδα:

Η πρωτογενής πρόληψη περιλαμβάνει τη μείωση της έκθεσης σε παράγοντες κινδύνου, μέσω της νομοθεσίας, της ρύθμισης, της εκπαίδευσης και της αλλαγής της συμπεριφοράς.

Η δευτερογενής πρόληψη περιλαμβάνει τον έλεγχο, την ανίχνευση και την ιατρική πρόληψη. Η αποφυγή περιττής θεραπείας και η βελτιωμένη σχέση κόστους-αποτελεσματικότητας του διαγνωστικού ελέγχου μπορεί να προέλθει από τον εντοπισμό υποομάδων υψηλού κινδύνου στον πληθυσμό, συμπεριλαμβανομένης της χρήσης βιολογικών δεικτών.

Η τριτογενής πρόληψη αποσκοπεί στην πρόληψη της υποτροπής.

Το ιδανικό βάρος είναι δείκτης υγείας. Ο έλεγχος και η διαχείριση του αποτελεί ένα σημαντικό μέρος της προληπτικής διαχείρισης της υγείας. Οι παρεμβάσεις δε στη διατροφή αποτελούν μέρος της διαχείρισης του ασθενούς και των ασθενειών τόσο σε ποιοτικό όσο και σε ποσοτικό επίπεδο.

5.2 Μοντελοποίηση

Η μοντελοποίηση της διαδικασίας αξιολόγησης του ατόμου σε βιοχημικό επίπεδο και όχι μόνο παρέχει ταχύτητα και ακρίβεια και μπορεί να αποτελέσει χρήσιμο εργαλείο για την αντιμετώπιση προβλημάτων σε πληθυσμιακό επίπεδο. Το κόστος και η προσβασιμότητα αποτελούν σημαντικό παράγοντα αποτελεσματικότητας, αλλά το σημαντικότερο παράγοντα αποτελεί η προσβασιμότητα σε πλήθος δεδομένων για την διαμόρφωση και εκπαίδευση συστημάτων τεχνητής νοημοσύνης υψηλής προβλεψιμότητας.

5.3 Αυτοματισμοί

Η χρησιμότητα των μοντέλων μπορεί να αποσβεστεί μέσω διασύνδεσης με μεγάλο μέρος του πληθυσμού μέσω διαδραστικών εργαλείων web και mobile

συσκευών. Τα συστήματα αυτά μπορούν να εμπλουτίζονται με νέα δεδομένα από τους χρήστες, και να δημιουργούν πρωτογενώς κατηγοριοποιήσεις και διαιτητικές προτάσεις χτίζοντας “awareness” για τη σημαντικότητα του ιδανικού βάρους και της σημαντικότητας της διατροφής.

5.4 Προοπτική

Η ερμηνεία και η εξαγωγή περιεχομένου από σύνθετα μεταβολικά σύνολα δεδομένων είναι εξαιρετικά απαιτητική και αντιπροσωπεύει έναν σημαντικό τομέα έρευνας. Ακόμα, σημειώνονται μεγάλα βήματα για την ενσωμάτωση των δεδομένων του μεταβολισμού με τη γονιδιωματική και την πρωτεωμική. Η Metabolomics προσφέρει την υπόσχεση ότι, στο μέλλον, η βιοχημική ανάλυση από τον γονότυπο στον φαινότυπο θα μετρηθεί και θα διερευνηθεί για νέες γνώσεις σχετικά με τη βιολογία του ανθρώπινου σώματος και την ιατρική.

Πηγές και βιβλιογραφία

1	Bradley A. Perkins et al 2017	Precision medicine screening using whole-genome sequencing and advanced imaging to identify disease risk in adults
2	Chandra Shekhar Pareek 2011	Sequencing technologies and genome sequencing
3	Christopher B Newgard ¹ , Jie An, James R Bain, Michael J Muehlbauer, Robert D Stevens, Lillian F Lien, Andrea M Haqq, Svati H Shah, Michelle Arlotto, Cris A Slentz, James Rochon, Dianne Gallup, Olga Ilkayeva, Brett R Wenner, William S Yancy Jr, Howard Eisenson, Gerald Musante, Richard S Surwit, David S Millington, Mark D Butler, Laura P Svetkey	A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance
4	D. Hammerl, M. Smid, A.M. Timmermans, S. Sleijfer, J.W.M. Martens, R. Debets	Breast cancer genomics and immuno-oncological markers to guide immune therapies
5	EC 10/3/2016	Regulatory framework applicable in the field of personalised medicine.
6	Elizabeth T. Cirulli, ^{1,8,*} Lining Guo, ² Christine Leon Swisher, ¹ Naisha Shah, ¹ Lei Huang, ¹ Lori A. Napier, ¹ Ewen F. Kirkness, ¹ Tim D. Spector, ³ C. Thomas Caskey, ⁴ Bernard Thorens, ⁵ J. Craig Venter, ⁶ and Amalio Telenti ^{7,**}	Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk
7	<u>J Hesketh</u>	Personalised nutrition: How far has nutrigenomics progressed?
8	<u>Maria Daniela Hurtado A</u> ¹ , <u>Andres Acosta</u> ²	Precision Medicine and Obesity
9	Nitish Nag, Vaibhav Pandey, Ramesh Jain	Live Personalized Nutrition Recommendation Engine
10	<u>Pedro González-Muniesa</u> ^{1,2,3,4} and <u>J. Alfredo Martínez</u> ^{1,2,3,4,5,*}	Precision Nutrition and Metabolic Syndrome Management
11	<u>Priscila Alves Maranhão</u> , ¹ <u>Gustavo Marísio Bacelar-Silva</u> , ¹ <u>Duarte Nuno Gonçalves Ferreira</u> , ¹ <u>Conceição Calhau</u> , ^{1,2} <u>Pedro Vieira-Marques</u> , ¹ and <u>Ricardo João Cruz-Correia</u> ¹	Nutrigenomic Information in the openEHR Data Set

12	<u>Roberto Martín-Hernández, Guillermo Reglero, José M Ordovás, Alberto Dávalos</u>	NutriGenomeDB: a nutrigenomics exploratory and analytical platform
13	Ruan Elliott, Teng Jin Ong	Nutritional genomics
14	<u>Sridhar A. Malkaram, Yousef I. Hassan, and Janos Zempleni</u>	Online Tools for Bioinformatics Analyses in Nutrition Sciences
15	<u>Tristan Cordier, Anders Lanzén, Laure Apothéoz-Perret-Gentil, Thorsten Stoeck</u>	Embracing Environmental Genomics and Machine Learning for Routine Biomonitoring
16	https://epic.iarc.fr/contactus/principalinvestigatorscentres.php	
17	https://gnomad.broadinstitute.org	
18	https://hmdb.ca/bmi_metabolomics	
19	https://hmdb.ca/bmi_metabolomics?page=12	
20	https://towardsdatascience.com/predicting-micronutrients-using-neural-networks-and-random-forest-part-1-83a1469766d7	
21	https://twinsuk.ac.uk/resources-for-researchers/access-our-data/	
22	https://www.cdc.gov/obesity/data/adult.html	
23	https://www.kaggle.com/cdc/national-health-and-nutrition-examination-survey	
24	https://www.ncbi.nlm.nih.gov/medgen/18127	
25	https://www.nutrigenetics.net/StartYourResearch.aspx	

Παράρτημα

Βιοχημικό προφίλ - Εύρος τιμών

Lab Tests and Values		
Test	Normal Values	
Albumin, serum	3.5 – 5.5 g/dL	
ALT	45 U/L or less	
AST	40 U/L or less	
Bilirubin, total	1.2 – 1.3 mg/dL or less	
Blood group (A,B,O)		
BUN, serum	7-25 mg/dL	
Calcium, serum	2.2-2.6 mmol/L	
CBC with differential	values given with report	
Chloride, serum	96-109 mmol/L	
CPK	male: 17 – 148 U/L, female: 10 – 70 U/L	
Creatinine, serum	0.6 – 6 mg/dL	
CRP	0.8 mg/dL or less	
Glucose, plasma	70-110 mg/dL	less than 140 mg/dL
LDH	less than 240 U/L	
Potassium, serum	3.5 – 5.3 mmol/L	
Prothrombin time	9 – 12 seconds	
Rapid plasma reagin	nonreactive	
Sedimentation rate	male: 0 – 15 mm/hr, female: 0 – 20 mm/hr	
Sodium, serum	135 – 147 mmol/L	
Triiodothyronine (T3)	85 – 205 mg/mL	
Thyroxine (T4)	4.5 – 12 mg/dL	
Total protein, serum	6 – 8.5 g/dL	
Triglycerides	less than 150 mg/dL	
Uric acid, serum	male: 3.9 – 9 mg/dL, female: 2.2 – 7.7 mg/dL	

Lab Tests and Values		
Urinalysis, routine		
Bicarbonate		
Clinical Use:		
Reference Ranges:	22 – 29 mmol/l	
Associated Diseases:	Higher than normal levels: Hyperaldosteronism Cushing syndrome Lower than normal levels: Ketoacidosis Lactic acidosis Kidney disease Addison disease	
Aspartate Aminotransferase		
	<ul style="list-style-type: none"> • Males: 10 to 40 units/L 	
	<ul style="list-style-type: none"> • Females: 9 to 32 units/L 	
<ul style="list-style-type: none"> • Serum globulin: 	<ul style="list-style-type: none"> • 2.0 to 3.5 grams per deciliter (g/dL) or 20 to 35 grams per liter (g/L) 	

Στοιχεία διαίτας

Variable Name	Variable Description
DR1STY	Did {you/SP} add any salt to {your/her/his} food at the table yesterday? Salt includes ordinary or seasoned salt, lite salt, or a salt substitute.
DR1TACAR	Alpha-carotene (mcg)
DR1TALCO	Alcohol (gm)
DR1TATOA	Added alpha-tocopherol (Vitamin E) (mg)
DR1TATOC	Vitamin E as alpha-tocopherol (mg)
DR1TB12A	Added vitamin B12 (mcg)

Variable Name	Variable Description
DR1TBCAR	Beta-carotene (mcg)
DR1TCAFF	Caffeine (mg)
DR1TCALC	Calcium (mg)
DR1TCARB	Carbohydrate (gm)
DR1TCHL	Total choline (mg)
DR1TCHOL	Cholesterol (mg)
DR1TCOPP	Copper (mg)
DR1TCRYP	Beta-cryptoxanthin (mcg)
DR1TFA	Folic acid (mcg)
DR1TFDFE	Folate as dietary folate equivalents (mcg)
DR1TFF	Food folate (mcg)
DR1TFIBE	Dietary fiber (gm)
DR1TFOLA	Total folate (mcg)
DR1TIRON	Iron (mg)
DR1TKCAL	Energy (kcal)
DR1TLYCO	Lycopene (mcg)

Variable Name	Variable Description
DR1TLZ	Lutein + zeaxanthin (mcg)
DR1TM161	MFA 16:1 (Hexadecenoic) (gm)
DR1TM181	MFA 18:1 (Octadecenoic) (gm)
DR1TM201	MFA 20:1 (Eicosenoic) (gm)
DR1TM221	MFA 22:1 (Docosenoic) (gm)
DR1TMAGN	Magnesium (mg)
DR1TMFAT	Total monounsaturated fatty acids (gm)
DR1TMOIS	Moisture (gm)
DR1TNIAC	Niacin (mg)
DR1TNUMF	Total number of foods/beverages reported in the individual foods file
DR1TP182	PFA 18:2 (Octadecadienoic) (gm)
DR1TP183	PFA 18:3 (Octadecatrienoic) (gm)
DR1TP184	PFA 18:4 (Octadecatetraenoic) (gm)
DR1TP204	PFA 20:4 (Eicosatetraenoic) (gm)
DR1TP205	PFA 20:5 (Eicosapentaenoic) (gm)
DR1TP225	PFA 22:5 (Docosapentaenoic) (gm)

Variable Name	Variable Description
DR1TP226	PFA 22:6 (Docosahexaenoic) (gm)
DR1TPFAT	Total polyunsaturated fatty acids (gm)
DR1TPHOS	Phosphorus (mg)
DR1TPOTA	Potassium (mg)
DR1TPROT	Protein (gm)
DR1TRET	Retinol (mcg)
DR1TS040	SFA 4:0 (Butanoic) (gm)
DR1TS060	SFA 6:0 (Hexanoic) (gm)
DR1TS080	SFA 8:0 (Octanoic) (gm)
DR1TS100	SFA 10:0 (Decanoic) (gm)
DR1TS120	SFA 12:0 (Dodecanoic) (gm)
DR1TS140	SFA 14:0 (Tetradecanoic) (gm)
DR1TS160	SFA 16:0 (Hexadecanoic) (gm)
DR1TS180	SFA 18:0 (Octadecanoic) (gm)
DR1TSELE	Selenium (mcg)
DR1TSFAT	Total saturated fatty acids (gm)

Variable Name	Variable Description
DR1TSODI	Sodium (mg)
DR1TSUGR	Total sugars (gm)
DR1TTFAT	Total fat (gm)
DR1TTHEO	Theobromine (mg)
DR1TVARA	Vitamin A as retinol activity equivalents (mcg)
DR1TVB1	Thiamin (Vitamin B1) (mg)
DR1TVB12	Vitamin B12 (mcg)
DR1TVB2	Riboflavin (Vitamin B2) (mg)
DR1TVB6	Vitamin B6 (mg)
DR1TVC	Vitamin C (mg)
DR1TVD	Vitamin D (D2 + D3) (mcg)
DR1TVK	Vitamin K (mcg)
DR1TWS	When you drink tap water, what is the main source of the tap water? Is the city water supply (community water supply); a well or rain cistern; a spring; or something else?
DR1TZINC	Zinc (mg)
DRQSDIET	Are you currently on any kind of diet, either to lose weight or for some other health-related reason?

Variable Name	Variable Description
DRQSPREP	How often is ordinary salt or seasoned salt added in cooking or preparing foods in your household? Is it never, rarely, occasionally, or very often?
DR1_300	Was the amount of food that {you/NAME} ate yesterday much more than usual, usual, or much less than usual?
DR1_320Z	Total plain water drank yesterday - including plain tap water, water from a drinking fountain, water from a water cooler, bottled water, and spring water.
DR1_330Z	Total tap water drank yesterday - including filtered tap water and water from a drinking fountain.
DR1BWATZ	Total bottled water drank yesterday (gm)

Standard Biochemistry Profile

Variable Name	Variable Description
LBDSALSI	Albumin (g/L)
LBDSBUSI	Blood urea nitrogen (mmol/L)
LBDS CASI	Total calcium (mmol/L)
LBDSCHSI	Cholesterol (mmol/L)
LBDS CRSI	Creatinine (umol/L)
LBDSGBSI	Globulin (g/L)
LBDSGLSI	Glucose, refrigerated serum (mmol/L)
LBDSIRSI	Iron, refrigerated serum (umol/L)
LBDSPHSI	Phosphorus (mmol/L)
LBDS TBSI	Total bilirubin (umol/L)
LBDS TPSI	Total protein (g/L)
LBDS TRSI	Triglycerides, refrigerated (mmol/L)
LBDSUASI	Uric acid (umol/L)
LBXSAL	Albumin (g/dL)
LBXSAPSI	Alkaline phosphatase (IU/L)
LBXSASSI	Aspartate aminotransferase AST (IU/L)

Variable Name	Variable Description
LBXSATSI	Alanine aminotransferase ALT (IU/L)
LBXSBU	Blood urea nitrogen (mg/dL)
LBXSC3SI	Bicarbonate (mmol/L)
LBXSCA	Total calcium (mg/dL)
LBXSCH	Cholesterol (mg/dL)
LBXSCK	Creatine Phosphokinase(CPK) (IU/L)
LBXSCLSI	Chloride (mmol/L)
LBXSCR	Creatinine (mg/dL)
LBXSGB	Globulin (g/dL)
LBXSGL	Glucose, refrigerated serum (mg/dL)
LBXSGTSI	Gamma glutamyl transferase (U/L)
LBXSIR	Iron, refrigerated serum (ug/dL)
LBXSKSI	Potassium (mmol/L)
LBXSLDSI	Lactate dehydrogenase (U/L)
LBXSNASI	Sodium (mmol/L)
LBXSOSI	Osmolality (mmol/Kg)

Variable Name	Variable Description
LBXSPH	Phosphorus (mg/dL)
LBXSTB	Total bilirubin (mg/dL)
LBXSTP	Total protein (g/dL)
LBXSTR	Triglycerides, refrigerated (mg/dL)
LBXSUA	Uric acid (mg/dL)
SEQN	Respondent sequence number.