



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ - ΙΑΤΡΙΚΗ ΣΧΟΛΗ

Πρόγραμμα Μεταπτυχιακών Σπουδών

Δημόσια Υγεία- Πρωτοβάθμια Φροντίδα Υγείας - Υπηρεσίες Υγείας

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Βαθμολογικές Κλίμακες Αξιολόγησης Βαρύτητας των Σοβαρά
Πασχόντων Ασθενών:**

**Διερεύνηση της προγνωστικής ακρίβειας και κλινικής
χρησιμότητας σε ασθενείς της Μονάδας Εντατικής
Θεραπείας του Πανεπιστημιακού Νοσοκομείου Ηρακλείου**

Ελισάβετ Γενειατάκη
Νοσηλεύτρια ΠΕ

Επιβλέπων:

Ευάγγελος Ι. Κριτωτάκης,
Αναπληρωτής Καθηγητής
Βιοστατιστικής

Συνεπιβλέπουσα:

Αικατερίνη Βαπορίδη, Αναπληρώτρια
Καθηγήτρια Ιατρικής Εντατικής
Θεραπείας

Ηράκλειο, [Φεβρουάριος 2022]

© 2022

Elisavet Geneiataki & Academic Supervisors

ALL RIGHTS RESERVED

Πίνακας περιεχομένων

1.1.Περίληψη	5
1.2. Abstract	8
2. Εισαγωγή	10
2.1. Η Μονάδα Εντατικής Θεραπείας (ΜΕΘ)	10
1.1.1. Ιστορικά Στοιχεία.....	10
1.1.2. Η εξέλιξη των Μονάδων Εντατικής Θεραπείας – Η εισαγωγή έρευνας στην εντατικολογία	10
1.2. Συστήματα Αξιολόγησης Βαρύτητας Νόσου- Έκβασης στη ΜΕΘ	11
1.2.1. Λόγοι δημιουργίας τους	11
1.2.2. Καταστάσεις που αξιολογούν οι Κλίμακες Βαρύτητας	12
1.2.3. Κατηγορίες Κλιμάκων Βαρύτητας.....	13
1.2.4 Χρήση Κλιμάκων Βαρύτητας στην καθημερινή Κλινική Πράξη.....	13
1.3. Κλίμακες που αναλύθηκαν στην παρούσα μελέτη	14
1.3.1. Κλίμακες Αξιολόγησης Βαρύτητας Νοσήματος κατά την εισαγωγή στη ΜΕΘ - APACHE (Acute Physiology And Chronic Health Evaluation)	15
2.3.2. Κλίμακες Αξιολόγησης Βαρύτητας Νοσήματος κατά την εισαγωγή στη ΜΕΘ - SAPS (Simplified Acute Physiology Score)	16
2.3.3. Κλίμακες Αξιολόγησης Δυσλειτουργίας ζωτικών οργάνων κατά την εισαγωγή στη ΜΕΘ - SOFA (Sequential Organ Failure Assessment)	17
2.3.4. Σύγκριση των κλιμάκων σύμφωνα με τα αποτελέσματα τους	17
1.4. Scores και αποφάσεις φροντίδας ασθενών	18
1.5. Στόχοι και Σημασία μελέτης	19
3. Μεθοδολογία	21
3.1. Ερευνητικό Σχέδιο	21
3.2. Συλλογή Δεδομένων	21
3.3. Κλίμακες Βαρύτητας	22
3.3.1. Πίνακας 1- Αναλυτική Περιγραφή των συλλεγόμενων δεδομένων	22
3.3.2. Πίνακας 2- Acute Physiologic and Chronic Health Evaluation (APACHE) II ¹	25
3.3.3. Πίνακας 3- Simplified Acute Physiologic Score (SAPS) II ¹	27
3.3.4. Πίνακας 4 -Sequential Organ Failure Assessment (SOFA) ¹	28
3.4. Προστασία Προσωπικών Δεδομένων	29
3.5. Μέγεθος δείγματος (Αριθμός ασθενών – Περίοδος μελέτης)	29
3.6. Στατιστική Ανάλυση	30
3.6.1 Περιγραφική Στατιστική	30
3.6.2 Διμεταβλητές Αναλύσεις	30
3.6.3 Αξιολόγηση Προγνωστικής Ακρίβειας και Κλινικής Χρησιμότητας	31

3.6.4. Μέτρα Διαχωριστικής Ικανότητας	31
3.6.5. Μέτρα Βαθμονόμησης	36
3.7. Έγκριση από Επιτροπή Βιοηθικής	45
4. Αποτελέσματα	46
4.1. Χαρακτηριστικά πληθυσμού μελέτης	46
4.2. Προγνωστική ακρίβεια της πρόβλεψης θανάτου στη ΜΕΘ	50
4.3.Καθαρό όφελος και ανάλυση καμπυλών απόφασης	55
4.4. Πρόβλεψη ενδονοσοκομειακής θνησιμότητας.....	58
4.5. Παράρτημα συμπληρωματικών διαγραμμάτων.....	61
5. Συζήτηση.....	63
6. Συμπέρασμα.....	67
7. Βιβλιογραφία	68

1.1.Περίληψη

**Τίτλος: Βαθμολογικές Κλίμακες Αξιολόγησης βαρύτητας των σοβαρά πασχόντων ασθενών:
Διερεύνηση της προγνωστικής ακρίβειας και κλινικής χρησιμότητας σε ασθενείς της Μονάδας
Εντατικής Θεραπείας του Πανεπιστημιακού Νοσοκομείου Ηρακλείου**

Από: Ελισάβετ Γενειατάκη, BSc, RN

Επιβλέποντες: Ευάγγελος Ι. Κριτωτάκης, Αναπληρωτής Καθηγητής Βιοστατιστικής

Αικατερίνη Βαπορίδη, Αναπληρώτρια Καθηγήτρια Ιατρικής Εντατικής Θεραπείας

Ημερομηνία: Φεβρουάριος 2022

Εισαγωγή: Η σωστή διαχείριση κρίσιμων νοσημάτων περιλαμβάνει τη γρήγορη και ακριβή ταυτοποίηση των βαρέως πασχόντων ασθενών. Αρκετά προγνωστικά μοντέλα (κλίμακες και βαθμολογίες βαρύτητας) έχουν προταθεί για αυτό το σκοπό- 3 από τα πιο ευρέως γνωστά και εφαρμοσμένα περιλαμβάνουν το Acute Physiology and Chronic Health Disease Classification System II (APACHE II), το Sequential Organ Failure Assessment (SOFA) και το Simplified Acute Physiology Score II (SAPS II). Τα προγνωστικά αυτά μοντέλα χρησιμοποιούν κλινικές και εργαστηριακές μεταβλητές, ώστε να προβλέψουν την ενδονοσοκομειακή θνησιμότητα και το αποτέλεσμα τους είναι μια εκτιμώμενη πιθανότητα του θανάτου εντός του νοσοκομείου, η οποία θεωρείται μια εκτίμηση της σοβαρότητας της νόσου. Στην εντατική θεραπεία, τα μοντέλα αυτά χρησιμοποιούνται ως επί το πλείστον για τη διαστρωμάτωση του κινδύνου (Benchmarking) κατά τη σύγκριση αποτελεσμάτων και φροντίδας σε διάφορες συνθήκες, για παράδειγμα ανάμεσα σε διαφορετικά νοσοκομεία ή στο ίδιο νοσοκομείο σε διαφορετικούς χρόνους. Ωστόσο, αυτά τα μοντέλα στερούνται ευρείας εξωτερικής επικύρωσης της ακρίβειας πρόβλεψής τους σε ποικίλες ομάδες βαρέως πασχόντων, συμπεριλαμβανομένου στο περιβάλλον μιας μικτής Μονάδας Εντατικής Θεραπείας (ΜΕΘ) ενηλίκων στην Ελλάδα. Για να είναι κλινικά χρήσιμο ένα προγνωστικό μοντέλο θα πρέπει να έχει καλή προγνωστική απόδοση (διάκριση και βαθμονόμηση) όταν δοκιμάζεται στις συνθήκες όπου θα χρησιμοποιηθεί, ώστε να αποδείξει τη χρησιμότητα του στη λήψη αποφάσεων.

Σκοπός: Η μελέτη αυτή είχε σαν στόχο την εκτέλεση εξωτερικής επικύρωσης, που περιλαμβάνει σύγκριση της πρόβλεψης του APACHE II, SOFA και SAPS II για κάθε ασθενή έναν προς έναν, για τη

θνησιμότητα εντός της ΜΕΘ και εντός του νοσοκομείου, σε μια ανεξάρτητη ομάδα από βαρέως πάσχοντες.

Μεθοδολογία: Πραγματοποιήθηκε μια αναδρομική μελέτη κοορτής από το Μάρτιο του 2018 έως το Φεβρουάριο του 2020, σε μια μικτή παθολογική- χειρουργική ΜΕΘ, αποτελούμενη από 12 κλίνες ενός Πανεπιστημιακού Νοσοκομείου στην Ελλάδα. Για να εκτιμηθεί η αποτελεσματικότητα των SAPS II, APACHE II και SOFA στην πρόβλεψη της θνησιμότητας εντός της ΜΕΘ και εντός του νοσοκομείου, εξετάστηκε η ικανότητα διάκρισης, με ανάλυση καμπύλης ROC και βαθμονόμηση μεταξύ της προβλεπόμενης θνησιμότητας, μέσω γραφημάτων και ελέγχου Hosmer- Lemeshow. Επιπλέον, εξετάστηκε αν τα μοντέλα αυτά μπορούν να βοηθήσουν στη λήψη αποφάσεων, μέσω Ανάλυσης Καμπύλης Απόφασης.

Αποτελέσματα: Έγινε ανάλυση δεδομένων από 778 ασθενείς (μέση ηλικία 68 έτη, 64% άνδρες). Οι μέσες τιμές των SAPS II, APACHE II και SOFA ήταν ήταν $40,9 \pm 13,0$, $23,6 \pm 7,4$ και $9,9 \pm 3,5$ αντίστοιχα. Η θνησιμότητα εντός ΜΕΘ ήταν 21,9% (95% διάστημα εμπιστοσύνης [CI] 19,0 - 24,9%) και η ενδονοσοκομειακή θνησιμότητα ήταν 34,7% (95% CI 31,4 - 38,2%). Για την πρόβλεψη της θνησιμότητας εντός ΜΕΘ, βρέθηκε από τα αποτελέσματα των δεικτών ότι μπορούν να διακρίνουν αποδεκτά και εξίσου καλά, με εμβαδόν καμπυλών ROC (AUC) 0,75 (95% CI 0,71-0,80), 0,72 (95% CI 0,67-0,76) και 0,71 (95% CI 0,67-0,76) για SAPS II, Apache II και SOFA αντίστοιχα ($p = 0,204$). Το SAPS II εμφανίζει εξαιρετική βαθμονόμηση σε ολόκληρο το φάσμα πιθανοτήτων θανάτου εντός της ΜΕΘ, αλλά το APACHE II και το SOFA είχαν κακή βαθμονόμηση σε ασθενείς υψηλού κινδύνου (το πρώτο υπερεκτιμώντας και το δεύτερο υποεκτιμώντας την πιθανότητα, όταν ο κίνδυνος υπερέβαινε το 40%). Για την πρόβλεψη της νοσοκομειακής θνησιμότητας, το SAPS II και το APACHE II είχαν σημαντικά καλύτερη ικανότητα διακρίσεων από το SOFA [AUCS 0,77 (95% CI 0,73-0,80), 0,76 (95% CI 0,72-0,79 και 0,70 (95% CI 0,66-0,74) αντίστοιχα, $p = 0,002$)] , ενώ η βαθμονόμηση μεταξύ παρατηρούμενων και προβλεπόμενων πιθανοτήτων ήταν καλή για τις τρεις βαθμολογίες. Η ανάλυση καμπύλης απόφασης έδειξε ότι, το SAPS II έχει αυξήσει το καθαρό κλινικό όφελος σε ένα ευρύ φάσμα «κατωφλιών» κινδύνου (από περίπου 15% έως 80%) για την θνησιμότητα εντός ΜΕΘ και εντός νοσοκομείου. Το APACHE II έδειξε χαμηλότερο καθαρό όφελος από το SAPS II για θνησιμότητα εντός της ΜΕΘ, αλλά συγκρίσιμο καθαρό όφελος για τη νοσοκομειακή θνησιμότητα. Το SOFA είχε χαμηλότερο καθαρό όφελος από το SAPS II και το APACHE II τόσο για τη θνησιμότητα εντός της ΜΕΘ όσο και για την

νοσοκομειακή θνησιμότητα, αλλά παρ'όλα αυτά υψηλότερο από τις εναλλακτικές στρατηγικές, όπως η χρήση της ηλικίας του ασθενούς για τη λήψη αποφάσεων.

Συμπέρασμα: Σε αυτή τη μελέτη παρατηρήθηκε ότι, οι προγνωστικές βαθμολογίες SAPS II, Apache II και SOFA είναι επαρκή εργαλεία για την πρόβλεψη θνησιμότητας εντός της ΜΕΘ και εντός του νοσοκομείου σε μια ομάδα ασθενών από μια μικτή ΜΕΘ στην Ελλάδα. Το SAPS II ήταν το μοντέλο καλύτερης απόδοσης, τόσο για τη θνησιμότητα εντός της ΜΕΘ όσο και για τη νοσοκομειακή θνησιμότητα, αλλά το APACHE II θα μπορούσε επίσης να είναι χρήσιμο για την αναγνώριση των ασθενών που κινδυνεύουν να πεθάνουν στο νοσοκομείο. Όλες οι βαθμολογίες, ειδικά το SAPS II, έδειξαν πιθανότητες να χρησιμοποιηθούν ως βοηθητικά εργαλεία για τη στήριξη της λήψης αποφάσεων από τους ιατρούς και τις οικογένειες σχετικά με το επίπεδο της θεραπευτικής διαχείρισης και της παρηγορητικής φροντίδας.

Λέξεις-κλειδιά: Προγνωστικές βαθμολογίες, πρόβλεψη θνησιμότητας, εντατική φροντίδα, Μονάδα Εντατικής Θεραπείας, μοντέλα υποστήριξης αποφάσεων.

1.2. Abstract

Title: Severity of illness scoring systems for critically ill patients: an investigation of prognostic accuracy and clinical utility in patients treated in the Intensive Care Unit of the University Hospital of Heraklion

By: Elisavet Geneiataki, BSc, RN

Supervisors: Evangelos I. Kritsotakis, Associate Professor of Biostatistics

Aikaterini Varoridi, Associate Professor in Intensive Care Medicine

Date: February 2022

Background: Managing critical illness requires quick and accurate identification of the critically ill patients. Several prognostic models (severity scales and scores) have been proposed for this aim – three of the most widely known and broadly applied include the Acute Physiology and Chronic Health Disease Classification System II (APACHE II), the Sequential Organ Failure Assessment (SOFA), and the Simplified Acute Physiology Score II (SAPS II). These prognostic models use clinical and laboratory variables to predict in-hospital mortality and their output is an estimated probability of in-hospital death, which is regarded as an estimate of disease severity. In critical care, such models are mostly utilized for risk stratification (benchmarking) when comparing outcomes and care across settings, for example between hospitals or of one hospital over time. However, these models are lacking wide external validation of predictive accuracy in broad cohorts of critically ill patients, including the setting of the mixed adult Intensive Care Unit (ICU) in Greece. To be clinically useful, a prognostic model should have good predictive performance (discrimination and calibration) when tested in the setting where it will be used and demonstrate usefulness for decision-making.

Objective: This study aimed to perform external validation involving head-to-head comparisons of the predictive performance of APACHE II, SOFA and SAPS II for intra-ICU and in-hospital mortality in an independent cohort of critically ill patients.

Methods: We conducted a retrospective cohort study from March 2018 to February 2020, in a 12-bed mixed medical-surgical ICU at a University Hospital in Greece. To assess the effectiveness of SAPS II,

APACHE II, and SOFA to predict intra-ICU and in-hospital mortality we examined discrimination ability through ROC curve analysis, and calibration between predicted probabilities and observed mortality rates via graphical methods and the Hosmer-Lemeshow goodness-of-fit test. Moreover, we examined whether these models may aid decision-making by means of Decision Curve Analysis.

Results: We analyzed data from 778 patients (median age 68 years, 64% male). Mean SAPS II, APACHE II, and SOFA scores were 40.9 ± 13.0 , 23.6 ± 7.4 , and 9.9 ± 3.5 respectively. Intra-ICU mortality was 21.9% (95% confidence interval [CI] 19.0 – 24.9%) and in-hospital mortality was 34.7% (95%CI 31.4 - 38.2%). For intra-ICU mortality prediction, index scores were found to discriminate acceptably and equally well, with areas under ROC curves (AUC) of 0.75 (95%CI 0.71-0.80), 0.72 (95%CI 0.67-0.76) and 0.71 (95%CI 0.67-0.76) for SAPS II, APACHE II, and SOFA respectively ($p = 0.204$). SAPS II displayed excellent calibration across the full range of probabilities for intra-ICU death, but APACHE II and SOFA had poor calibration in high-risk patients (the former over-estimated and latter under-estimated risks exceeding 40%). For in-hospital mortality prediction, SAPS II and APACHE II had significantly better discrimination ability than SOFA (AUCs 0.77 (95%CI 0.73-0.80), 0.76 (95%CI 0.72-0.79), and 0.70 (95%CI 0.66-0.74) respectively; $p = 0.002$), while calibration between observed and predicted probabilities was good for all three scores. Decision curve analysis showed that SAPS II has increased net clinical benefit across a wide range of risk thresholds (from about 15% to 80%) for intra-ICU and in-hospital mortality. APACHE II displayed lower net benefit than SAPS II for intra-ICU mortality but comparable net benefit for in-hospital mortality. SOFA had lower net benefit than SAPS II and APACHE II for both intra-ICU and in-hospital mortality, but nevertheless higher than alternative strategies such as using patient's age for decision making.

Conclusion: In this study, we observed that SAPS II, APACHE II, and SOFA prognostic scores are adequate tools for intra-ICU and in-hospital mortality prediction in a cohort of patients from a mixed adult ICU in Greece. The SAPS II model was the best performing model for both intra-ICU and in-hospital mortality, but APACHE II could also be helpful for identifying patients at risk of dying in the hospital. All scores, especially SAPS II, exhibited potential to be used as ancillary tools to support decision-making by physicians and families regarding the level of therapeutic management and palliative care.

Key words: Prognostic scores, mortality prediction, critical care, intensive care unit, decision support models.

2. Εισαγωγή

2.1. Η Μονάδα Εντατικής Θεραπείας (ΜΕΘ)

1.1.1. Ιστορικά Στοιχεία

Η Μονάδα Εντατικής Θεραπείας είναι το ειδικά σχεδιασμένο και εξειδικευμένο τμήμα του νοσοκομείου για την παρακολούθηση και αντιμετώπιση επειγουσών και απειλητικών για τη ζωή καταστάσεων, οι οποίες όμως είναι δυνητικά αναστρέψιμες και απαιτούν πολύπλευρη φροντίδα¹. Σύμφωνα με ιστορικά στοιχεία, κατά τη διάρκεια του πολέμου της Κριμαίας, το 1854, η Florence Nightingale εφάρμοσε για πρώτη φορά τη διαλογή τραυματισμένων στρατιωτών σύμφωνα με τη βαρύτητα τους, τοποθετώντας εκείνους που έχρηζαν εντατικότερης νοσηλευτικής φροντίδας πλησιέστερα στο νοσηλευτικό σταθμό². Αυτό θεωρήθηκε η πρώτη προσομοίωση μιας σύγχρονης Μονάδας Εντατικής Θεραπείας. Παίρνοντας στοιχεία από τέτοιου είδους προσπάθειες, οι οποίες βασιζόνταν σε κοινό τρόπο οργάνωσης και στόχο, ξεκίνησαν να εμφανίζονται οι πρώτες οργανωμένες ΜΕΘ, με σκοπό την ανάνηψη βαρέως πασχόντων ασθενών, τη δεκαετία του 1950.

1.1.2. Η εξέλιξη των Μονάδων Εντατικής Θεραπείας – Η εισαγωγή έρευνας στην εντατικολογία

Από την εμφάνιση της πρώτης ΜΕΘ, με την πρόοδο του τομέα της ιατρικής και της τεχνολογίας, η φροντίδα των ασθενών που χρήζουν εντατική νοσηλεία έχει εξελιχθεί σημαντικά. Πρόκειται πλέον για ένα από τα ταχύτερα αναπτυσσόμενα τμήματα στο χώρο ενός νοσοκομείου².

Οι αρχικές Μονάδες Εντατικής Θεραπείας στις ΗΠΑ χωρίζονταν μάλιστα σύμφωνα με την φύση των περιστατικών, σε χειρουργικές, παθολογικές, αναπνευστικές, καρδιολογικές και νευροχειρουργικές¹. Πέραν των εξειδικευμένων γιατρών, οι περισσότεροι των οποίων κατέχουν και ειδικότητα εντατικολογίας, στελεχώνεται και από εξίσου εξειδικευμένους νοσηλευτές, καθώς και πολλούς άλλους επαγγελματίες υγείας².

Στις μέρες μας, η προσέγγιση των ασθενών, ιδίως των βαρέως πασχόντων τείνει να μετατρέπεται από εμπειρική σε τεκμηριωμένη και βασισμένη σε ενδείξεις, από όλα τα επαγγέλματα υγείας. Πλέον, δίνεται λιγότερο βάρος στη διαίσθηση και στη μη συστηματική κλινική εμπειρία, και περισσότερη προσοχή συγκεντρώνει η συστηματική αξιολόγηση τεκμηρίων κλινικών μελετών, επιτυγχάνοντας έτσι

τον βέλτιστο συνδυασμό¹. Ο απώτερος σκοπός όλων αυτών είναι καταρχάς η μείωση της νοσηρότητας και θνησιμότητας και κατ' επέκταση η ελαχιστοποίηση της διάρκειας παραμονής στη ΜΕΘ και της χρήσης ιατροφαρμακευτικών πόρων³.

1.2. Συστήματα Αξιολόγησης Βαρύτητας Νόσου- Έκβασης στη ΜΕΘ

1.2.1. Λόγοι δημιουργίας τους

Οι ραγδαίες τεχνολογικές εξελίξεις στον τομέα της σύγχρονης εντατικολογίας έχουν οδηγήσει στην νοσηλεία όλο και πιο πολύπλοκων περιστατικών, στην αναζήτηση και μελέτη νέων παρεμβατικών και φαρμακευτικών μεθόδων, συνεπώς και στην αποτελεσματικότητα και την απόδοση τους, καθώς και στην πρόγνωση και έκβαση των νοσημάτων. Συνάμα, έχουν υπάρξει εξίσου σημαντικές θεσμικές και οικονομικές αλλαγές, οι οποίες έχουν συμβάλει στον τρόπο λειτουργίας των Μονάδων Εντατικής Θεραπείας³. Για το λόγο αυτό έκαναν την αρχική τους εμφάνιση τα Συστήματα Αξιολόγησης βαρύτητας Νόσου ή Έκβασης².

Τα συστήματα αυτά εμφανίστηκαν και εξελίχθηκαν τα τελευταία 25 με 30 χρόνια. Δημιουργήθηκαν κυρίως για τους βαρέως πάσχοντες που νοσηλεύονταν στις Μονάδες Εντατικής Θεραπείας, με σκοπό την αξιολόγηση της σοβαρότητας της νόσου τους, την πρόγνωση της έκβασης τους και την πρόβλεψη του κινδύνου θανάτου^{1,2}. Η αρχική συσχέτιση της βαρύτητας με την έκβαση προσδιορίστηκε με μια αναδρομική συλλογή δεδομένων, από διάφορες κατηγορίες ασθενών, όπως παθολογικά περιστατικά, χειρουργικά περιστατικά, κλπ. και ταυτόχρονη καταγραφή της έκβασης τους. Από τα αποτελέσματα, δημιουργήθηκαν κάποια αρχικά πρότυπα συστημάτων, τα οποία αξιολογήθηκαν προοπτικά σε μεγαλύτερους πληθυσμούς ασθενών. Καθώς άλλαζαν τα χαρακτηριστικά των ασθενών και οι πρακτικές της εντατικής θεραπείας με την πάροδο του χρόνου, τα συστήματα εκτίμησης εξελίσσονταν και αξιολογούνταν διαρκώς, έως και σήμερα².

Ένας από τους βασικούς σκοπούς δημιουργίας συστημάτων αξιολόγησης βαρύτητας νόσου είναι η σωστή επιλογή περιστατικών που χρειάζονται νοσηλεία στη ΜΕΘ και ο στόχος να γίνει περισσότερο αποτελεσματική μετά την εισαγωγή τους, μέσω παροχής σημαντικών πληροφοριών, απαραίτητων για τον στρατηγικό σχεδιασμό της. Συμπληρωματικά, ένα από τα κύρια σημεία ήταν η προσπάθεια παροχής ακόμη πιο ποιοτικής φροντίδας στην πλειονότητα των ασθενών, με πιο στοχευμένες υπηρεσίες³. Σε γενικές λοιπόν γραμμές, τα συστήματα αυτά συμβάλουν στη λήψη κλινικών

αποφάσεων, συμμετέχουν στην έρευνα και βοηθούν στη σύγκριση της ποιότητας σχετικά με την ιατρική και νοσηλευτική φροντίδα μεταξύ διαφορετικών ΜΕΘ¹.

1.2.2. Καταστάσεις που αξιολογούν οι Κλίμακες Βαρύτητας

Ένα από τα αρχικά συστήματα αξιολόγησης βαρύτητας της νόσου ήταν το APGAR, το οποίο καθιερώθηκε το 1953, με στόχο την εκτίμηση της ζωτικότητας ενός νεογέννητου². Ακολούθησαν μεταγενέστερα εξίσου σημαντικά και ευρέως χρησιμοποιούμενα ως και σήμερα η Κλίμακα Γλασκώβης, με σκοπό τον προσδιορισμού του επιπέδου συνείδησης ενός ασθενή³.

Τις τελευταίες τρεις δεκαετίες έχει αναπτυχθεί ένας μεγάλος αριθμός συστημάτων αξιολόγησης σχετικά με τη βαρύτητα και την πρόγνωση της έκβασης για τα περιστατικά των Μονάδων Εντατικής Θεραπείας¹. Η πλειοψηφία των συστημάτων αυτών αποτελούνται από κλίμακες που δίνουν στο χρήστη μια τελική αριθμητική τιμή (ή βαθμολογία βαρύτητας), σύμφωνα με κάποιες μεταβλητές που τις αποτελούν. Η τελική αυτή τιμή ποσοτικοποιεί τελικά τη βαρύτητα της νόσου, την πιθανότητα θανάτου, αλλά και τη διάρκεια νοσηλείας που πρόκειται να ακολουθήσει¹. Οι κλίμακες αυτές έχουν φτωχή κλινική σημασία σε ασθενείς που νοσηλεύονται εκτός Μονάδας Εντατικής Θεραπείας¹.

Οι κύριες εφαρμογές τους είναι επομένως η αξιολόγηση βαρύτητας της κατάστασης υγείας των ασθενών της ΜΕΘ, η πρόβλεψη κινδύνου θανάτου, η αξιολόγηση της έκτασης που τυχόν έχει πάρει μια πολυοργανική δυσλειτουργία, η εκτίμηση ποιότητας της παρεχόμενης φροντίδας και η σύγκριση της μεταξύ διαφορετικών ΜΕΘ ή της ίδιας σε διαφορετικές χρονικές περιόδους. Ακόμη, συμβάλουν στην αξιολόγηση της βαρύτητας ανάμεσα σε διαφορετικές ομάδες βαρέως πασχόντων και έτσι, στην επιλογή και τη διαστρωμάτωση των ασθενών που συμμετέχουν σε κλινικές δοκιμές, αλλά και στο σχεδιασμό βέλτιστης διαχείρισης των διαθέσιμων πόρων².

Η λήψη θεραπευτικών αποφάσεων όμως με βάση τα συστήματα αυτά είναι αμφιλεγόμενη, καθότι ο σχεδιασμός τους βασίζεται σε ομάδες ασθενών, και όχι μεμονωμένων περιπτώσεων. Επιπλέον, ένας σημαντικός περιορισμός τους αποτελεί η αδυναμία πρόβλεψης σχετικά με την ποιότητα ζωής μακροπρόθεσμα².

1.2.3. Κατηγορίες Κλιμάκων Βαρύτητας

Οι κλίμακες βαρύτητας χωρίζονται σε κατηγορίες σύμφωνα με το τι εκτιμούν. Μια κατηγορία είναι εκείνες που αξιολογούν τη βαρύτητα του νοσήματος κατά την εισαγωγή του ασθενή στη ΜΕΘ και συγχρόνως εκτιμούν και τον κίνδυνο θανάτου, δίνοντας μια σημαντική εικόνα σχετικά με την έκβαση^{1,2}. Δεύτερη κατηγορία αποτελούν οι κλίμακες που επικεντρώνονται στη δυσλειτουργία ζωτικών οργάνων, τόσο κατά την εισαγωγή του ασθενή, όσο και σε όλη τη διάρκεια νοσηλείας². Τα πιο γνωστά σκορ που ανήκουν στην πρώτη κατηγορία είναι το APACHE (Acute Physiology And Chronic Health Evaluation), το SAPS (Simplified Acute Physiology Score) και το MPM (Mortality Probability Model)¹. Αντίστοιχα στη δεύτερη συμπεριλαμβάνονται το MODS (Multiple Organ Dysfunction Score), το SOFA (Sequential Organ Failure Assessment) και το LODS (Logistic Organ Dysfunction Score)³.

Με το πέρασμα των χρόνων, και καθώς εξελίσσονται τα χαρακτηριστικά του πληθυσμού, οι επιπτώσεις των νοσημάτων και οι παρεμβάσεις στη φροντίδα, και λόγω των εκσυγχρονισμένων στατιστικών τεχνικών, τα συστήματα αυτά εξίσου αλλάζουν και προσαρμόζονται σε νεότερες εκδόσεις, ώστε να διασφαλίζεται συνεχώς η ακρίβεια των μετρήσεων³.

Τα συστήματα ασφαλώς έχουν και ορισμένους περιορισμούς. Από τους πιο σημαντικούς είναι ότι, οι κλίμακες που σχετίζονται με την εκτίμηση της οργανικής δυσλειτουργίας δεν έχουν ικανοποιητική ακρίβεια στην πρόβλεψη της επιβίωσης. Επίσης, δεν πρέπει να παραλείπεται το γεγονός ότι, η βαρύτητα ενός νοσήματος αλλά και η επηρεασμένη λειτουργία οργάνων διαφέρουν μεταξύ των ασθενών, όσο και στον ίδιο τον ασθενή σε διαφορετικές χρονικές στιγμές, κάτι το οποίο ισχύει και για τις δύο κατηγορίες κλιμάκων³.

2.2.4 Χρήση Κλιμάκων Βαρύτητας στην καθημερινή Κλινική Πράξη

Ασφαλώς, η δημιουργία των δεικτών βαρύτητας και η διεξοδική μελέτη τους πραγματοποιείται με σκοπό την συμβολή τους στην λήψη αποφάσεων. Είναι σημαντικό κατά τη χρήση τους σε θέματα καθημερινής κλινικής πράξης να λαμβάνονται υπόψιν ορισμένα σημαντικά σημεία. Η ΜΕΘ αποτελεί ένα κατάλληλο περιβάλλον για την εφαρμογή προγνωστικών συστημάτων, λόγω του ότι ο πληθυσμός των ασθενών είναι συγκεκριμένος κάθε χρονική στιγμή, η φροντίδα σαφώς καθορισμένη και η σοβαρότητα του νοσήματος είναι στενά συνδεδεμένη με την ενδονοσοκομειακή θνητότητα¹.

Αρχικά, η κλίμακα θα πρέπει να σχετίζεται με μια σημαντική κλινική έκβαση, όπως για τη μελέτη της πιθανότητας ενδονοσοκομειακού θανάτου¹. Επίσης, χρειάζεται να είναι εύκολη στην πρόσβαση και στη χρήση από το προσωπικό που χρησιμοποιείται, καθώς η συλλογή των δεδομένων που συμβάλουν στο αποτέλεσμα συνήθως είναι χρονοβόρα και δαπανηρή¹.

Με τον ορισμό «ισχύς» (validity) περιγράφεται η αξιολόγηση της απόδοσης ενός μοντέλου πρόβλεψης, με εφαρμογή μιας κλίμακας σε ένα σύνολο δεδομένων. Γίνεται λοιπόν σαφές ότι, και οι κλίμακες που αναφέρονται θα πρέπει να έχουν αρκετή ισχύ και διάκριση (discrimination), όπως να είναι και καλά βαθμονομημένες¹. Στο σημείο αυτό, δεν πρέπει να παραληφθεί και η σημασία της ακρίβειας της κλίμακας. Όσο περισσότερο η ομάδα ασθενών που μελετάται ταιριάζει με εκείνη επάνω στην οποία αναπτύχθηκε η κλίμακα, τόσο αυξάνεται η ακρίβεια².

Όπως αναφέρθηκε και παραπάνω, οι δείκτες/κλίμακες βαρύτητας παρέχουν μια καλή ένδειξη του κινδύνου θανάτου και της σοβαρότητας κατάστασης υγείας σε ομάδες ασθενών ΜΕΘ, αλλά αυτό δεν ισχύει απαραίτητα για εξατομικευμένη πρόγνωση¹ ή για την αξιολόγηση μιας πάθησης ξεχωριστά, πχ τη μελέτη της προβλεπόμενης θνητότητας ασθενών με ηπατική ανεπάρκεια, AIDS, κλπ¹. Επισημαίνεται επίσης ότι η ακρίβεια των αποτελεσμάτων είναι φτωχή για ασθενείς εκτός του περιβάλλοντος της ΜΕΘ¹.

Κρίσιμης σημασίας για σωστά και αξιόπιστα αποτελέσματα των κλιμάκων είναι η συνεχής ενημέρωση τους. Γενικώς όλα τα προγνωστικά συστήματα βαρύτητας πρέπει να ενημερώνονται σε τακτά χρονικά διαστήματα, με βάση τα σύγχρονα δεδομένα. Διαφορετικά, υπάρχει ο κίνδυνος να μην συμπεριληφθούν σημαντικές επιπτώσεις νέων μεθόδων και τεχνολογιών στη φροντίδα. Συνεπακόλουθα, αυτό ενδέχεται να οδηγήσει σε σταδιακή απώλεια της ικανότητας τους να προβλέπουν με ακρίβεια ορισμένες εκβάσεις, συμπεριλαμβανομένου και του ενδονοσοκομειακού θανάτου¹. Είναι λοιπόν απαραίτητη η συνεχής επαναξιολόγηση και αναβάθμιση τους, ώστε να συμβαδίζουν με τη σύγχρονη κλινική πρακτική και τα στοιχεία των ασθενών την εκάστοτε χρονική περίοδο¹.

1.3. Κλίμακες που αναλύθηκαν στην παρούσα μελέτη

Στην εργασία αυτή επιλέχθηκαν να εξεταστούν τρεις κλίμακες, οι οποίες χρησιμοποιούνται ευρέως στη ΜΕΘ και διαρκώς εκσυγχρονίζονται, ως ακολούθως.

1.3.1. Κλίμακες Αξιολόγησης Βαρύτητας Νοσήματος κατά την εισαγωγή στη ΜΕΘ - APACHE (Acute Physiology And Chronic Health Evaluation)

Η κλίμακα APACHE αναπτύχθηκε για πρώτη φορά το 1981 σε ασθενείς ΜΕΘ στις ΗΠΑ, με σκοπό την ταξινόμηση τους ανάλογα με τη βαρύτητα του νοσήματος². Στην πορεία, αναθεωρήθηκε το 1985, στο APACHE II, μορφή η οποία είναι από τα πιο ευρέως χρησιμοποιούμενα συστήματα βαρύτητας έως και σήμερα². Ασφαλώς έχουν προκύψει και πιο πρόσφατες εκδόσεις, η APACHE III και IV. Με τα μοντέλα αυτά αναπτύχθηκαν εξισώσεις για την πρόβλεψη διάρκειας παραμονής ενός ασθενούς στη ΜΕΘ⁴. Καθώς για την κατασκευή των κλιμάκων αυτών χρησιμοποιήθηκε μια μεγαλύτερη βάση δεδομένων από διάφορα νοσοκομεία, τα παλαιότερα μοντέλα αναδιαμορφώθηκαν, περιέχοντας διαφορετικές μεταβλητές πρόβλεψης⁴. Εκτός από την πρόγνωση για τη διάρκεια παραμονής, μπορούν να χρησιμοποιηθούν και ως δείκτες αναφοράς για την αξιολόγηση της φροντίδας και για σύγκριση της αποδοτικότητας των πόρων της ΜΕΘ⁴. Τα τελευταία δύο μοντέλα αυτά όμως είναι λιγότερο εύχρηστα, λόγω των πολλών παραμέτρων που περιλαμβάνουν².

Η κλίμακα APACHE II βασίζεται σε μια σειρά κλινικών στοιχείων από τα οποία προκύπτει μια βαθμολογία βαρύτητας, που χρησιμοποιείται σε μια εξίσωση λογιστικής παλινδρόμησης, η οποία τελικά προβλέπει την ενδονοσοκομειακή θνητότητα¹. Οι τιμές που επιλέγονται είναι οι μη φυσιολογικές τιμές που παρατηρήθηκαν κατά τη διάρκεια του πρώτου 24ωρου από την στιγμή εισαγωγής του ασθενούς στη ΜΕΘ^{1,2}.

Η κλίμακα APACHE II απαρτίζεται από τρία μέρη. Στο πρώτο, αθροίζεται η βαθμολογία 12 κλινικοεργαστηριακών παραμέτρων, λαμβάνοντας υπόψιν τις πιο παθολογικές τιμές των πρώτων 24 ωρών νοσηλείας στη ΜΕΘ. Στο δεύτερο μέρος βαθμολογείται η ηλικία και στο τρίτο η γενική κατάσταση υγείας του ασθενή πριν την εισαγωγή, σε συνάρτηση με την αιτία εισόδου³. Η τιμή του αθροίσματος αντικατοπτρίζει τον κίνδυνο θανάτου. Έτσι, μια τιμή APACHE II score που ισούται με 25 αντιστοιχεί σε προβλεπόμενη θνητότητα 50%, ενώ τιμές >30 αντιστοιχούν σε θνητότητα 80%³. Η μέγιστη τιμή που μπορεί να πάρει ένα σκορ είναι το 71³.

Όλες οι κλίμακες APACHE έχουν εξαιρετική ακρίβεια πρόβλεψης ποσοστών θνησιμότητας στις ΜΕΘ. Όπως και τα περισσότερα μοντέλα, απαιτεί περιοδική επαναξιολόγηση, αναθεώρηση και αναβάθμιση, καθώς οι νέες θεραπευτικές παρεμβάσεις επηρεάζουν έντονα τη θνητότητα, κάνοντας τα σκορ να

μεταβάλλονται συνεχώς¹. Επίσης, για την APACHE II ισχύει το γεγονός ότι, η έκβαση των ασθενών ενδέχεται να μην μπορεί να προβλεφθεί με ακρίβεια για συγκεκριμένες υποομάδες (πχ σε εκείνους με ηπατικές ανεπάρκειες, σήψεις, μαιευτικές νόσους, AIDS κλπ). Αυτό οφείλεται στο γεγονός ότι, παρ'όλο που τα προγνωστικά συστήματα βαρύτητας προέρχονται από μεγάλες βάσεις δεδομένων, αποδεικνύονται πολύ μικρές για να επιτρέψουν να αξιολογηθεί κάθε πάθηση ξεχωριστά. Έτσι, η προβλεπόμενη θνητότητα ασθενών για ορισμένες από αυτές μπορεί να είναι ανακριβής¹.

2.3.2. Κλίμακες Αξιολόγησης Βαρύτητας Νοσήματος κατά την εισαγωγή στη ΜΕΘ - SAPS (Simplified Acute Physiology Score)

Η κλίμακα SAPS (Simplified Acute Physiologic Score) αναπτύχθηκε και σταθμίστηκε στη Γαλλία το 1984¹. Το χαρακτηριστικό της ήταν η χρήση ηλικίας και άλλων 13 σταθμισμένων μεταβλητών, με σκοπό την πρόβλεψη κινδύνου θανάτου σε ασθενείς της ΜΕΘ^{1,5}. Όπως και στην κλίμακα APACHE, έτσι και στη συγκεκριμένη, λαμβάνονται υπόψιν για τον υπολογισμό της οι χειρότερες τιμές κατά τη διάρκεια των πρώτων 24 ωρών από την εισαγωγή του ασθενούς στη ΜΕΘ¹.

Το 1993 έγινε αναβάθμιση της αρχικής κλίμακας, μέσω λογιστικής παλινδρόμησης και έτσι αναπτύχθηκε η SAPS II, η πλέον περισσότερο ευρέως χρησιμοποιούμενη μορφή². Η SAPS II περιλαμβάνει 17 μεταβλητές, 12 μεταβλητές σχετικά με την κατάσταση υγείας, την ηλικία του ασθενή, την αιτία εισαγωγής και την ύπαρξη ή όχι συγκεκριμένων χρόνιων νοσημάτων (μεταστατικού καρκίνου, αιματολογικής κακοήθειας, AIDS). Εκτός από την τελευταία παράμετρο, η οποία είτε υπάρχει είτε όχι, όλες οι υπόλοιπες κατηγοριοποιήθηκαν σύμφωνα με το εύρος των τιμών, με την κάθε κατηγορία να αντιστοιχεί σε συγκεκριμένους βαθμούς².

Όσο υψηλότερη είναι η τιμή της SAPS II για έναν ασθενή, τόσο πιο βαριά είναι η κατάσταση του και τόσο αυξάνεται ο κίνδυνος ενδοσοκομειακής θνητότητας¹. Σε γενικές γραμμές, η SAPS II παρουσιάζει εξαιρετική ακρίβεια πρόβλεψης για ποσοστά θνητότητας¹.

Η συνέχεια αυτών ήταν η κλίμακα SAPS III, η οποία κυκλοφόρησε το 2005. Η συγκεκριμένη περιλαμβάνει 20 μεταβλητές, οι οποίες χωρίζονται σε τρεις υποκατηγορίες, τα χαρακτηριστικά του ασθενή πριν την εισαγωγή, την αιτία εισαγωγής και το βαθμό διαταραχής των τιμών που καταγράφονται εντός μιας ώρας από την εισαγωγή (σε αντίθεση με το χρονικό παράθυρο των 24 ωρών του SAPS II)¹. Οι μεταβλητές προσαρμόζονται για την πρόγνωση της νοσοκομειακής

θνησιμότητας σε επτά γεωγραφικές περιοχές και κυμαίνεται από 0 έως 217 βαθμούς¹. Ακόμη, φαίνεται να έχει καλή διακριτική ικανότητα, βαθμονόμηση και προσαρμογή⁵⁻⁷.

2.3.3. Κλίμακες Αξιολόγησης Δυσλειτουργίας ζωτικών οργάνων κατά την εισαγωγή στη ΜΕΘ - SOFA (Sequential Organ Failure Assessment)

Το SOFA score δημιουργήθηκε από ομάδα ειδικών επιστημόνων το 1994. Τα αρχικά του γράμματα αρχικά συμβόλιζαν «Sepsis-related Organ Assessment», καθώς χρησιμοποιούνταν για να περιγράψει ποσοτικά την οργανική δυσλειτουργία σε σηπτικούς ασθενείς της ΜΕΘ. Στην πορεία διαπιστώθηκε ότι εφαρμόζεται εξίσου καλά και σε μη σηπτικούς ασθενείς, οπότε το ακρωνύμιο του άρχισε να αναφέρεται στο «Sequential Organ Failure Assessment», που ισχύει έως και σήμερα⁷⁻⁹.

Πλέον χρησιμεύει στην εκτίμηση της λειτουργίας 6 οργανικών συστημάτων ή ζωτικών οργάνων. Αρχικά, του αναπνευστικού, με το λόγο PaO₂/ FiO₂, του κυκλοφορικού, από τη δοσολογία αγγειοσυσπαστικών φαρμάκων που χορηγούνται στον ασθενή για τη σταθεροποίηση του, του ήπατος, από την τιμή χολερυθρίνης ορού, του αιμοποιητικού από τον αριθμό αιμοπεταλίων, του Κεντρικού νευρικού, από την τιμή της Κλίμακας Γλασκώβης (GCS) και των νεφρών, από την τιμή κρεατινίνης ορού και το ημερήσιο ποσό αποβαλλόμενων ούρων⁷⁻⁹. Το γεγονός ότι στην τελική τιμή συμπεριλαμβάνεται η δοσολογία φαρμάκων αποτελεί ένα επισφαλές στοιχείο, καθώς τα θεραπευτικά πρωτόκολλα ποικίλουν, ωστόσο είναι δύσκολο να αποφευχθεί⁷⁻⁹.

Κάθε λειτουργία βαθμολογείται αναλόγως από 0 έως 4 και το άθροισμα τους δίνει την τελική τιμή του SOFA, με εύρος τιμών από 0 έως 24⁷. Οι τιμές που λαμβάνονται υπόψιν προέρχονται από το πρώτο 24ωρο μετά την εισαγωγή. Η τιμή αυτή είναι ικανοποιητική στην πρόβλεψη της θνητότητας και η μεταβολή της κατά τη διάρκεια της νοσηλείας έχει επίσης προγνωστική αξία². Για παράδειγμα, η αύξηση του score κατά τις πρώτες 48 ώρες νοσηλείας στη ΜΕΘ συνδυάζεται με θνητότητα τουλάχιστον 50%².

2.3.4. Σύγκριση των κλιμάκων σύμφωνα με τα αποτελέσματα τους

Είναι σαφές ότι, οι περισσότερες κλίμακες βαρύτητας στη ΜΕΘ, από τις αρχικές τους μορφές έως και οι πιο πρόσφατες εκδοχές τους έχουν στο σύνολο τους ακρίβεια που θεωρείται αποδεκτή και

αντιπροσωπευτική, τόσο για τη νοσηρότητα όσο και για τη θνητότητα. Ωστόσο, υπάρχουν αρκετές διαφορές και περιορισμοί μεταξύ τους¹.

Αρχικά, οι διαφορές εντοπίζονται σχετικά με την διαδικασία συλλογής δεδομένων. Ο τύπος των μεταβλητών που καταγράφονται και η στιγμή καταγραφής των μετρήσεων ποικίλει ανάμεσα στις κλίμακες. Για παράδειγμα, η κλίμακα APACHE περιλαμβάνει ένα ευρύ φάσμα στοιχείων, από την κλινική κατάσταση του ασθενή, έως τη γενική κατάσταση της υγείας του, ενώ οι μεταβλητές των υπολοίπων είναι περισσότερο συνοπτικές και πιο εύκολα μετρήσιμες.

Ο τρόπος με τον οποίο προκύπτει επίσης ο κίνδυνος θανάτου για κάθε μια κλίμακα διαφέρει, καθότι χρησιμοποιείται διαφορετική στατιστική και μαθηματική ανάλυση. Το κοινό σίγουρα που έχουν όλες είναι ότι, το τελικό score προέρχεται από το άθροισμα των επιμέρους παραμέτρων που αποτελείται η κάθε μεταβλητή.

Όσον αφορά την ικανότητα πρόβλεψης και τη σύγκριση μεταξύ τους, διάφορες μελέτες έχουν εξετάσει την απόδοση τους σε ομάδες περιστατικών. Μια συγκεκριμένη συστηματική ανασκόπηση που μελέτησε τις SOFA, SAPS II, APACHE II και APACHE III διαπίστωσε ότι, οι κλίμακες APACHE είναι κατά ένα ποσοστό ανώτερες στην πρόβλεψη θνητότητας στη ΜΕΘ από τις SOFA και SAPS¹.

Επίσης, σχετικά με το κόστος τους οι APACHE, SAPS και SOFA είναι εύκολα προσβάσιμες, χωρίς να απαιτούν ακριβό λογισμικό και οι δύο τελευταίες μάλιστα απαιτούν και μικρότερη συλλογή δεδομένων.

Η πλειοψηφία όμως των κλιμάκων, ασχέτως των διαφορών τους, έχουν σαν κοινό σημείο τον αρχικό στόχο δημιουργίας τους, την αξιολόγηση βαρύτητας και τον καθορισμό πρόγνωσης και στη συνέχεια αποδείχθηκαν χρήσιμες και αλλού. Κυρίως, διευκολύνουν την αξιολόγηση παρεμβάσεων, συγκρίνοντας ασθενείς παρόμοιου κινδύνου, την ποιότητα φροντίδας, τις εκβάσεις διαφορετικών ΜΕΘ εντός του ίδιου νοσοκομείου ή διαφορετικών και στη καλύτερη διαχείριση των υλικών και ανθρώπινων πόρων.

1.4. Scores και αποφάσεις φροντίδας ασθενών

Όμως, είναι γνωστό ότι οι βαθμολογικές κλίμακες βαρύτητας μπορεί να υποεκτιμούν ή να υπερεκτιμούν τη θνησιμότητα όταν χρησιμοποιούνται για συγκεκριμένες ομάδες ασθενών (π.χ.

καρκινοπαθείς), ειδικά αν αυτές οι υποομάδες δεν ήταν καλά αντιπροσωπευμένες στην αρχική κοόρτη¹⁰. Γενικά, η ακρίβεια των προγνωστικών μοντέλων μπορεί να αμφισβητηθεί όταν αυτά εφαρμόζονται σε διαφορετική γεωγραφική περιοχή ή όταν το φάσμα και η διαχείριση της νόσου μπορεί να ποικίλλουν ή να μεταβάλλονται με την πάροδο του χρόνου.

Κατά συνέπεια, πολλοί ερευνητές έχουν προσφάτως στρέψει το ενδιαφέρον τους στην εξωτερική επικύρωση (external validation) της προγνωστικής ακρίβειας των κλίμακων αξιολόγησης της βαρύτητας σοβαρά πασχόντων ασθενών σε τοπικά περιβάλλοντα περιθαλψής λίγο ή περισσότερο διαφορετικά από εκείνα στα οποία αναπτυχθήκαν και ελέγχθηκαν οι κλίμακες αυτές¹¹⁻¹⁵.

Επίσης προσφάτως, ερευνητές έχουν χρησιμοποιήσει καινοτόμες στατιστικές μεθόδους για να εξετάσουν το όφελος από την ενδεχομένη χρήση των κλίμακων αξιολόγησης της βαρύτητας για την υποστήριξη αποφάσεων που αφορούν τη διαχείριση ασθενών τελικού στάδιου και την παρηγορητική φροντίδα στις ΜΕΘ^{11,16}.

Δυο σημαντικές δυσκολίες που συναντώνται στη λήψη αποφάσεων σε καταστάσεις τέλους ζωής τους είναι η έλλειψη αξιόπιστης προγνωστικής βαθμολογικής κλίμακας και το γεγονός ότι η γνώμη των ασθενών λαμβάνεται πολύ σπανία υπόψη¹⁶. Αρκετές διεθνείς μελέτες έχουν δείξει μεγάλη διακύμανση στις αποφάσεις διαχείρισης των ασθενών τελικού στάδιου, η οποία έχει παρατηρηθεί και στις ελληνικές ΜΕΘ¹⁷. Η Ανάλυση Καμπυλών Απόφασης (Decision Curve Analysis - DCA) είναι στατιστική μέθοδος που αναπτύχθηκε για να αξιολογήσει και να συγκρίνει τα προγνωστικά μοντέλα και έχει το πλεονέκτημα ότι μπορεί να λάβει υπόψη τις επιθυμίες των ασθενών (ή των οικειών τους) όσο αφορά το βαθμό που είναι διατεθειμένοι να εκτεθούν στον κίνδυνο ψευδούς θετικής ή ψευδούς αρνητικής πρόγνωσης^{16,18,19}.

1.5. Στόχοι και Σημασία μελέτης

Η εργασία αυτή έχει στόχο να διερευνήσει την προγνωστική ακρίβεια των τεσσάρων κλίμακων βαρύτητας, APACHE II, SAPS II και SOFA, στο περιβάλλον της μικτής ΜΕΘ ενός μεγάλου ελληνικού νοσοκομείου. Επίσης, θα εξετάσει και θα συγκρίνει την κλινική χρησιμότητα (το “καθαρό όφελος” στην DCA) σε ενδεχόμενη χρήση αυτών προγνωστικών δεικτών για την υποστήριξη αποφάσεων που αφορούν την παρηγορητική φροντίδα ασθενών τελικού στάδιου στη ΜΕΘ.

Πιο συγκεκριμένα, ο σκοπός της DCA είναι να αξιολογήσει εάν η χρήση του προγνωστικού μοντέλου κατά μέσο ορό θα αποφέρει περισσότερα οφέλη παρά βλάβες. Η «πιθανότητα κατωφλιού» (threshold probability) χρησιμοποιείται ως μέτρο της βλάβης από άσκοπη θεραπεία (ψευδώς θετικά) σε σύγκριση με το όφελος της θεραπείας σε εκείνους που πραγματικά τη χρειάζονται (ορθώς θετικά). Το «καθαρό όφελος» (net benefit) αντιπροσωπεύει το όφελος της θεραπείας των σωστών ασθενών μείον τη βλάβη από τη θεραπεία εκείνων που δεν την χρειάζονται και υπολογίζεται ως σταθμισμένη διαφορά μεταξύ των ποσοστών των ορθώς θετικών και των ψευδώς θετικών^{16,18,19}. Τα αποτελέσματα παρουσιάζονται σε ένα γράφημα με τις πιθανότητες κατωφλιού επί της τετμημένης και το καθαρό όφελος του αξιολογημένου μοντέλου επί της τεταμένης. Επομένως, οι παραπάνω μέθοδοι αξιολόγησης και ανάλυσης μπορούν να συμβάλουν σημαντικά στη λήψη κρίσιμων αποφάσεων σχετικά με την αντιμετώπιση και τη θεραπεία των ασθενών.

3. Μεθοδολογία

3.1. Ερευνητικό Σχέδιο

Πρόκειται για αναδρομική μελέτη κοόρτης διαδοχικών ασθενών που νοσηλευτήκαν στη ΜΕΘ του Πανεπιστημιακού Γενικού Νοσοκομείου Ηράκλειου κατά τη διετία Μάρτιος 2018 έως Φεβρουάριος 2020 (registry-based retrospective cohort study). Τα απαιτούμενα στοιχεία συλλέχθηκαν από υπάρχουσες βάσεις δεδομένων της ΜΕΘ και του νοσοκομείου (CRITIS και MedLine/AMED, αντίστοιχα).

3.2. Συλλογή Δεδομένων

Οι πληροφορίες που συλλέχθηκαν για κάθε ασθενή για τους σκοπούς της μελέτης παρατίθενται αναλυτικά στον Πίνακα 1. Επιγραμματικά, για κάθε ασθενή ΜΕΘ εντός της διετούς περιόδου μελέτης οι πληροφορίες αυτές αφορούσαν:

1. Ιστορικό του ασθενούς: ημερομηνία εισαγωγής στο νοσοκομείο, ημερομηνία εισαγωγής στη ΜΕΘ, ηλικία, φύλο, ταξινόμηση περιστατικού, αίτια εισαγωγής/ICD-10, ιστορικό χρόνιου νοσήματος.
2. Κλινικο-εργαστηριακές παραμέτρους κατά τις πρώτες 24 ώρες από την εισαγωγή στη ΜΕΘ (χειρότερες τιμές) που είναι απαραίτητες για τον υπολογισμό των βαθμολογικών κλιμάκων APACHE II, SAPS II, LODS και SOFA, συμπεριλαμβανομένων των: Θερμοκρασία, Μέση αρτηριακή πίεση, Καρδιακός ρυθμός, Αναπνευστικός ρυθμός PaO₂/FiO₂, Αρτηριακό p, HCO₃⁻ ορού, Νάτριο ορού, Κάλιο ορού, Κρεατινίνη ορού, Αιματοκρίτης, Λευκά αιμοσφαίρια, Κλίμακα Κώματος Γλασκώβης GCS, Σφίξεις, Ούρα 24ώρου, Ουρία, Διττανθρακικά, Χολερυθρίνη, Αιμοπετάλια – Πηκτικότητα, Αιμοδυναμική κατάσταση.
3. Έκβαση του ασθενούς στη ΜΕΘ και το νοσοκομείο (θάνατος ή εξιτήριο, με την αντίστοιχη ημερομηνία).

Όλα τα παραπάνω στοιχεία αναζητήθηκαν στη βάση δεδομένων CRITIS της ΜΕΘ, εκτός από την έκβαση της νοσηλείας που αναζητήθηκε στη βάση MedLine/AMED του ΠαΓΝΗ (πίνακας 1). Επίσης, παρατίθενται οι κλίμακες που μελετήθηκαν στα πλαίσια της παρούσας έρευνας (Πίνακας 2-4).

3.3. Κλίμακες Βαρύτητας

3.3.1. Πίνακας 1- Αναλυτική Περιγραφή των συλλεγόμενων δεδομένων

ΠΕΡΙΓΡΑΦΗ ΠΑΡΑΜΕΤΡΟΥ/ΜΕΤΑΒΛΗΤΗΣ	ΤΙΜΕΣ (ενδεικτικά)	ΧΡΗΣΗ ΣΕ ΚΛΙΜΑΚΑ	ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ
Ιστορικό			
Κωδικοποιημένος (ανωνυμοποιημένος) αναγνωριστικός αριθμός ασθενούς	Αριθμητική τιμή		CRITIS, MedLine/Amed
Ημερομηνία εισαγωγής στο νοσοκομείο (πριν την εισαγωγή στη ΜΕΘ)	Ημερομηνία		
Ημερομηνία εισαγωγής στη ΜΕΘ	Ημερομηνία		
Ηλικία (σε έτη) την ημέρα εισαγωγής στη ΜΕΘ	Αριθμητική τιμή	Apache II, SAPS II	CRITIS
Φύλο	Α/Γ		CRITIS
Ταξινόμηση περιστατικού	Παθολογικό/Επείγον χειρουργείο/ Τακτικό χειρουργείο/ Τραύμα	SAPS II	CRITIS
Προέλευση εισαγωγής στη ΜΕΘ	Τμήμα Επειγόντων Περιστατικών /ΜΑΦ/Χειρουργείο/Κλινική		CRITIS
Αιτία εισαγωγής στη ΜΕΘ	ICD-10		CRITIS
Αν υπάρχει ιστορικό σοβαρής οργανικής ανεπάρκειας ή ανοσοκαταστολής:	Χωρίς τέτοιο ιστορικό/ Μη χειρουργημένος ασθενής/ Ασθενής που χειρουργήθηκε σε επείγουσα βάση/ Μετεγχειρητικός ασθενής μετά από προγραμματισμένη εγχείρηση	Apache II	CRITIS
Χρόνια Νόσος	Μεταστατικός καρκίνος / Αιματολογική Κακοήθεια/AIDS	SAPS II	CRITIS

Κατά τις πρώτες 24 ώρες από την εισαγωγή στη ΜΕΘ (χειρότερες τιμές):			
Θερμοκρασία (°C)	Αριθμητική τιμή	Apache II, SAPS II	CRITIS
Μέση αρτηριακή πίεση (mm Hg)	Αριθμητική τιμή	Apache II, SAPS II	CRITIS
Καρδιακός ρυθμός (κοιλιακή ανταπόκριση)	Αριθμητική τιμή	Apache II	CRITIS
Αναπνευστικός ρυθμός (μηχανικός ή μη αερισμός)	Αριθμητική τιμή	Apache II	CRITIS
Οξυγόνωση, FiO ₂ > 0,5: υπολογισμός AaDO ₂ (mm Hg)	Αριθμητική τιμή	Apache II	CRITIS
Οξυγόνωση, FiO ₂ < 0,5: καταγραφή μόνο PaO ₂ (mm Hg)	Αριθμητική τιμή	Apache II	CRITIS
Αρτηριακό pH (επί απουσίας αερίων αρτηριακού αίματος, καταγραφή HCO ₃ ⁻ ορού παρακάτω)	Αριθμητική τιμή	Apache II	CRITIS
HCO ₃ ⁻ ορού (φλεβικό αίμα – mmol/L) χρήση μόνο επί απουσίας αερίων αρτηριακού αίματος	Αριθμητική τιμή	Apache II	CRITIS
Νάτριο ορού (mmol/L or mEq/L)	Αριθμητική τιμή	Apache II, SAPS II	CRITIS
Κάλιο ορού (mmol/L or mEq/L)	Αριθμητική τιμή	Apache II, SAPS II	CRITIS
Κρεατινίνη ορού (mg/dL)	Αριθμητική τιμή	Apache II, SOFA	CRITIS
Αιματοκρίτης (%)	Αριθμητική τιμή	Apache II	CRITIS
Λευκά αιμοσφαίρια (1.000/mm ³)	Αριθμητική τιμή	Apache II, SAPS II	CRITIS
Κλίμακα Κώματος Γλασκώβης GCS (Βαθμολογία = 15 - Κλίμακα Γλασκώβης)	Αριθμητική τιμή	Apache II, SAPS II, SOFA	CRITIS
Σφίξεις (παλμοί/λεπτό)	Αριθμητική τιμή	SAPS II	CRITIS

PaO ₂ /FiO ₂ (mm Hg) αν σε μηχανική αναπνοή	Αριθμητική τιμή	SAPS II, SOFA	CRITIS
Ούρα 24ώρου (ml)	Αριθμητική τιμή	SAPS II	CRITIS
Ουρία (g/L)	Αριθμητική τιμή	SAPS II	CRITIS
Διττανθρακικά (mEq/L)	Αριθμητική τιμή	SAPS II	CRITIS
Χολερυθρίνη - Ήπαρ (mg/L)	Αριθμητική τιμή	SAPS II, SOFA	CRITIS
Αιμοπετάλια - Πηκτικότητα (x 1000/mm ³)	Αριθμητική τιμή	SOFA	CRITIS
Αιμοδυναμική κατάσταση	5 κατηγορίες με βάση Αρτ Πίεση, Dopamine, Adrenaline, Noradrenaline	SOFA	CRITIS
Έκβαση			
Έκβαση στη ΜΕΘ	Εξιτήριο / θάνατος		CRITIS
Ημερομηνία έκβασης στη ΜΕΘ	Ημερομηνία		CRITIS
Επανεισαγωγή στη ΜΕΘ (κατά την ίδια νοσηλεία)	Ναι / Όχι/ Άγνωστο		CRITIS
Ημερομηνία επανεισαγωγής στη ΜΕΘ	Ημερομηνία		CRITIS
Έκβαση στο Νοσοκομείο (κατά την ίδια νοσηλεία)	Εξιτήριο / μεταφορά σε άλλο νοσοκομείο / θάνατος		MedLine/Amed
Ημερομηνία έκβασης στο Νοσοκομείο	Ημερομηνία		MedLine/Amed

3.3.2. Πίνακας 2- Acute Physiologic and Chronic Health Evaluation (APACHE) II¹

ΦΥΣΙΟΛΟΓΙΚΗ ΠΑΡΑΜΕΤΡΟΣ	ΥΨΗΛΕΣ ΜΗ ΦΥΣΙΟΛΟΓΙΚΕΣ ΤΙΜΕΣ				ΦΥΣΙΟΛΟΓΙΚΕΣ ΤΙΜΕΣ	ΧΑΜΗΛΕΣ ΜΗ ΦΥΣΙΟΛΟΓΙΚΕΣ ΤΙΜΕΣ			
	+4	+3	+2	+1	0	+1	+2	+3	+4
1.Θερμοκρασία ορθού (°C)	>41	39-40,9		38,5-38,9	36-38,4	34-35,9	32-33,9	30-31,9	>29,9
2.Μέση Αρτηριακή πίεση (mmHg)	>160	130-159	110-129		70-109		50-69		<49
3.Καρδιακός ρυθμός (κοιλιακή ανταπόκριση)	>180	140-179	110-139		70-109		55-69	40-54	<39
4.Αναπνευστικός Ρυθμός (μηχανικός ή μη μη αερισμός)	>50	35-49		25-34	12-24	10-11	6-9		<5
5.Οξυγόνωση: AaDO ₂ ή PaO ₂ (mmHg)									
A) FiO ₂ >0,5: υπολογισμός AaDO ₂	>500	350-499	200-349		<200				
B) FiO ₂ <0,5: καταγραφή μόνο PaO ₂					>70	61-70		55-60	<55
6.Αρτηριακό pH (επί απουσίας αερίων αρτηριακού αίματος, καταγραφή HCO ₃ -ορού παρακάτω)	>7,7	7,6-7,69		7,5-7,59	7,33-7,49		7,25-7,32	7,15-7,24	<7,15
7.Νάτριο ορού (mmol/L)	>180	160-179	155-159	150-154	130-139		120-129	110-119	<110
8.Κάλιο ορού (mmol/L)	>7	6-6,9		5,6-5,9	3,5-5,4	3-3,4	2,5-2,9		<2,5
9.Κρεατίνη ορού (mg/dL) (διπλάσια βαθμολογία επί νεφρικής ανεπάρκειας)	>3,5	2-3,4	1,5-1,9		0,6-1,4		<0,6		
10.Αιματοκρίτης (%)	60		50-59,9	46-49,9	30-45,9		20-29,9		<20
11.Λευκά Αιμοσφαίρια (1000/ mm ³)	>40		20-39,9	15-19,9	3-14,9		1-2,9		<1
12.Κλίμακα Γλασκώβης GCS (Βαθμολογία=15-Κλίμακα Γλασκώβης)	15-GCS=								
ΟΛΙΚΗ ΒΑΘΜΟΛΟΓΙΑ (APS)	Άθροιση της βαθμολογίας των παραμέτρων 1-12=								

HCO ₃ - ορού (φλεβικό αίμα, mmol/L) Δεν προτιμάται μόνο επί απουσίας αερίων αρτηριακού αίματος	52	41-51,9	32-40,9	22-31,9	18-21,9	15-17,9
--	----	---------	---------	---------	---------	---------

- **Κατηγορία Ηλικιών σε έτη / Βαθμολογία αντίστοιχα** (<44/ 0, 44-54/2, 55-64/3, 65-74/ 5, >75/ 6)
- **Βαθμολογία Χρόνιας Νόσου**
 Αν ο ασθενής έχει ιστορικό σοβαρής οργανικής ανεπάρκειας ή είναι ανοσοκατεσταλμένος, αποδίδεται βαθμολογία σύμφωνα με τα παρακάτω:
 α. Μη χειρουργημένος ασθενής ή ασθενής που χειρουργήθηκε σε επείγουσα βάση: +5 βαθμοί
 β. Μετεγχειρητικός ασθενής μετά από προγραμματισμένη εγχείρηση: +2 βαθμοί
- **Βαθμολογία κατά APACHE II: APS+ Ηλικία + Χρόνια Νόσος**

3.3.3. Πίνακας 3- Simplified Acute Physiologic Score (SAPS) II¹

Simplified Acute Physiologic Score (SAPS II)						
ΒΑΘΜΟΛΟΓΙΑ = 12 ΜΕΤΑΒΛΗΤΕΣ+ ΗΛΙΚΙΑ+ ΧΡΟΝΙΑ ΝΟΣΟΣ+ ΕΙΔΟΣ ΕΙΣΑΓΩΓΗΣ						
Παράμετρος	Τιμή [Βαθμολογία]					
Σφίξεις (παλμοί/λεπτό)	<40 [11]	40-69 [2]	70-119 [0]	120-159 [4]	≥160 [7]	-
Συστολική Αρτηριακή Πίεση (mmHg)	<70 [13]	70-99 [5]	100-199 [0]	≥200 [2]	-	-
Θερμοκρασία (οC)	<39 [0]	≥39 [3]	-	-	-	-
ΡαΟ ₂ /FιΟ ₂ (mmHg)	<100 [11]	100-199 [9]	≥200 [6]	-	-	-
Ούρα 24ωρου (ml)	<500 [11]	500-999 [4]	≥1000 [0]	-	-	-
Ουρία (g/L)	<0,6 [0]	0,6-1,79 [6]	>1,8 [10]	-	-	-
Λευκά (*10 ³ /mm ³)	<1000 [12]	1000-19000 [0]	≥ 20000 [3]	-	-	-
Κάλιο (mEq/L)	<3 [3]	3-4,9 [0]	≥ 5 [2]	-	-	-
Νάτριο (mEq/L)	<125 [5]	125-144 [0]	≥ 145 [1]	-	-	-
Διττανθρακικά (mEq/L)	<15 [6]	15-19 [3]	>20 [0]	-	-	-
Χολερυθρίνη (mg/L)	<4 0	4-5,9 [4]	≥ 6 [9]	-	-	-
Κλίμακα Γλασκώβης	<6 [26]	6-8 [13]	9-10 [7]	11-13 [5]	14-15 [0]	-
Ηλικία (έτη)	<40 [0]	40-59 [7]	60-69 [12]	70-74 [15]	75-79 [16]	>80 [18]
Χρόνια Νόσος	Μεταστατικός Καρκίνος [9]	Αιματολογική Κακοήθεια [10]	AIDS [17]	-	-	-
Είδος Εισαγωγής	Προγραμματισμένη Χειρουργική επέμβαση [0]	Επείγουσα Χειρουργική επέμβαση [8]	Μη Χειρουργικό περιστατικό [6]	-	-	-
Βαθμολογία	29	40	52	64	77	-
Θνητότητα	10%	25%	50%	75%	90%	-

3.3.4. Πίνακας 4 -Sequential Organ Failure Assessment (SOFA)¹

Sequential Organ Failure Assessment (SOFA)						
Οργανικό Σύστημα	Παράμετρος	Βαθμολογία				
		4	3	2	1	0
Αναπνευστικό	PaO ₂ / FiO ₂ (mm Hg)	<100 και αναπνευστική υποστήριξη	<200 και αναπνευστική υποστήριξη	<300	<400	>400
Πηκτικότητα	Αιμοπετάλια (* 10 ³ /mm ³)	<20	<50	<100	<150	>150
Ήπαρ	Χολερυθρίνη (mg/dL)	>12	6-11,9	2-5,9	1,2-1,9	<1,2
Καρδιαγγειακό	Αιμοδυναμική Κατάσταση	Dopamine >15* Adrenaline >0,1* Noradrenaline >0,1*	Dopamine >5* Adrenaline ≤ 0,1* Noradrenaline ≤ 0,1*	Dopamine ≤ 5* Dobutamine όποια δόση	Μέση αρτηριακή πίεση (mmHg)<70	Μέση αρτηριακή πίεση (mmHg) >70
Κεντρικό Νευρικό	Κλίμακα Γλασκώβης	<6	6-9	10-12	13-14	15
Νεφροί	Κρεατινίνη (mg/dL) ή Διούρηση	>5 ή <500 (ml/d)	3,5 – 4,9 ή <200 (ml/d)	2-3,4	1,2-1,9	<1,2
ΒΑΘΜΟΛΟΓΙΑ		0-6	7-9	10-12	13-14	15-24
ΘΝΗΤΟΤΗΤΑ		<10%	15-20%	40-50%	50-60%	>80%

*αδρενεργικοί παράγοντες που χορηγούνται για τουλάχιστον 1 ώρα (οι δόσεις είναι σε mcg/min)

3.4. Προστασία Προσωπικών Δεδομένων

Η μελέτη ακολούθησε τους διεθνείς κανονισμούς για την προστασία προσωπικών δεδομένων (GDPR) καθ' όλη τη διάρκεια της διεξαγωγής, της ανάλυσης και της αναφοράς των αποτελεσμάτων της. Τονίζεται ότι, τα δεδομένα που συγκεντρώθηκαν για τους σκοπούς της μελέτης δεν περιλαμβάνουν πληροφορίες που θα μπορούσαν να ταυτοποιήσουν τους ασθενείς (όπως το ονοματεπώνυμο ή η ημερομηνία γέννησης). Επειδή ήταν σημαντικό να αποφευχθούν διπλοεγγραφές, να μπορούν να συνδεθούν στοιχεία του ίδιου ασθενούς μεταξύ των βάσεων CRITIS και AMED, και να μπορούν να γίνουν διορθώσεις σε δεύτερο χρόνο, αποθηκεύτηκε ένας κρυπτογραφημένος αριθμός αναγνώρισης του ασθενούς, ο οποίος θα συνδέεται (με αποκρυπτογράφηση) με μοναδικό αναγνωριστικό αριθμό που χρησιμοποιεί το νοσοκομείο (αριθμός περιστατικού). Ο αλγόριθμος αποκρυπτογράφησης είναι γνωστός μόνο στους ερευνητές της μελέτης.

3.5. Μέγεθος δείγματος (Αριθμός ασθενών – Περίοδος μελέτης)

Δεν υπάρχουν γενικές φόρμουλες υπολογισμού του ελάχιστου απαιτούμενου μεγέθους δείγματος σε μελέτες εξωτερικής επικύρωσης προγνωστικών μοντέλων. Ένας συχνά παρατιθέμενος κανόνας (rule of thumb) είναι ότι για μια έγκυρη αξιολόγηση της προγνωστικής ακρίβειας απαιτούνται κατ' ελάχιστο 100 συμβάντα για κάθε μελετώμενη έκβαση (π.χ. >100 θάνατοι εντός ΜΕΘ)¹¹. Ωστόσο, πρόσφατες μελέτες στατιστικής προσομοίωσης έδειξαν ότι πολύ συχνά ο αναγκαίος αριθμός συμβάντων πρέπει να υπερβαίνει τα 200 συμβάντα¹².

Λαμβάνοντας υπόψη ότι η θνησιμότητα στη ΜΕΘ κυμαίνεται (υπό συνηθισμένες συνθήκες) σε περίπου 25% με περί τις 400 νοσηλείες ετησίως, υπολογίσαμε ότι θα χρειαστούμε 2 έτη καταγραφής (περί τους 800 ασθενείς) ώστε να συμπεριλάβουμε περί τα 200 περιστατικά ενδονοσοκομειών θανάτων στη μελέτη μας. Για να μην επηρεαστεί η μελέτη από τις ειδικές συνθήκες της πανδημίας SARS-CoV-2, συλλέξαμε δεδομένα πριν τον Μάρτιο του 2020, δηλαδή κατά τη διετία Μάρτιος 2018 έως Φεβρουάριος 2020.

3.6. Στατιστική Ανάλυση

Η προγνωστική ακρίβεια των υπό μελέτη κλιμάκων βαρύτητας αξιολογήθηκε με στατιστικά μέτρα διακριτικής ικανότητας (discrimination) και βαθμονόμησης (calibration), συμπεριλαμβανομένων των καμπυλών ROC και των γραφημάτων βαθμονόμησης, που υπολογίστηκαν με το μοντέλο της πολλαπλής λογιστικής παλινδρόμησης²⁰. Χρησιμοποιήθηκε επίσης ανάλυση καμπυλών απόφασης (DCA) για να υπολογιστεί το καθαρό όφελος από την ενδεχόμενη κλινική χρήση των βαθμολογικών κλιμάκων βαρύτητας για την υποστήριξη αποφάσεων παρηγορητικής φροντίδας.

Η κωδικοποίηση και οργάνωση των δεδομένων της μελέτης έγιναν από την συγγραφέα με χρήση του λογισμικού Microsoft Excel. Το πρωτόκολλο της στατιστικής ανάλυσης δημιουργήθηκε από τον Αναπληρωτή Καθηγητή Βιοστατιστικής κ. Ευάγγελο Ι. Κριτσωτάκη και υλοποιήθηκε υπό την επίβλεψή του. Η πλειοψηφία των στατιστικών αναλύσεων έγιναν με το στατιστικό λογισμικό Stata, v.17.

3.6.1 Περιγραφική Στατιστική

Το πρώτο μέρος της ανάλυσης των δεδομένων περιέλαβε περιγραφική στατιστική ανάλυση, κατά την οποία έγινε σύνοψη των δημογραφικών και κλινικών χαρακτηριστικών, των προγνωστικών σκορς (SAPS II, Apache II, SOFA) και των εκβάσεων των ασθενών της μελέτης για κάθε μία μεταβλητή ξεχωριστά από τις υπόλοιπες. Για τα ποιοτικά χαρακτηριστικά (κατηγορικές και δυαδικές μεταβλητές), έγινε υπολογισμός των κατανομών των απόλυτων συχνοτήτων (αριθμός ατόμων, n) και των ποσοστιαίων συχνοτήτων. Για τις ποσοτικές μεταβλητές, εξετάστηκαν ιστογράμματα συχνοτήτων, τα οποία επικαλύφθηκαν με καμπύλες κανονικών κατανομών αντίστοιχης μέσης τιμής και τυπικής απόκλισης.

Στις περιπτώσεις που υπήρξε εμφανής ασυμμετρία συγκριτικά με την κανονική κατανομή, πραγματοποιήθηκε σύνοψη των ποσοτικών δεδομένων, χρησιμοποιώντας τη διάμεσο και το ενδοτεταρτημοριακό εύρος. Σε προσεγγιστικά συμμετρικές κατανομές, έγινε σύνοψη των δεδομένων με τη μέση τιμή \pm τυπική απόκλιση.

3.6.2 Διμεταβλητές Αναλύσεις

Πραγματοποιήθηκαν διμεταβλητές αναλύσεις (συγκρίσεις δύο ομάδων) για να εξεταστούν οι διαφορές των δημογραφικών και κλινικών χαρακτηριστικών και των προγνωστικών σκορς ανάμεσα σε ασθενείς που πέθαναν κατά τη νοσηλεία τους στη ΜΕΘ και ασθενείς που επιβίωσαν. Οι αναλύσεις αυτές

πραγματοποιήθηκαν και σε σχέση με την έκβαση στο νοσοκομείο. Η στατιστική σημαντικότητα των διαφορών σε ποιοτικές μεταβλητές (συγκρίσεις ποσοστών) ελέγχθηκε με τον μη-παραμετρικό έλεγχο χ^2 του Pearson. Για τις συγκρίσεις ποσοτικών (συνεχών) μεταβλητών, χρησιμοποιήθηκαν ιστογράμματα και θηκογράμματα για να διερευνηθούν οι μορφές των κατανομών. Για τη σύγκριση ποσοτικών δεδομένων με εμφανώς ασύμμετρες κατανομές ποσοτικών δεδομένων χρησιμοποιήθηκε μη-παραμετρική δοκιμασία αθροίσματος διατάξεων του Wilcoxon, ενώ για συμμετρικές κατανομές έγινε έλεγχος t ανεξάρτητων δειγμάτων.

3.6.3 Αξιολόγηση Προγνωστικής Ακρίβειας και Κλινικής Χρησιμότητας

Αξιολογήθηκε η προγνωστική ακρίβεια των υπό μελέτη προγνωστικών συστημάτων βαθμολόγησης βαρύτητας (SAPS II, Apache II, SOFA) σε ασθενείς ΜΕΘ χρησιμοποιώντας τόσο μέτρα διακριτικής/διαχωριστικής ικανότητας (discrimination) όσο και μέτρα βαθμονόμησης (calibration), συμπεριλαμβανομένων των καμπυλών ROC και των γραφημάτων βαθμονόμησης, που υπολογίστηκε με το στατιστικό μοντέλο λογιστικής παλινδρόμησης.

Χρησιμοποιήθηκε επίσης ανάλυση καμπυλών απόφασης (Decision Curve Analysis - DCA) για να υπολογιστεί το καθαρό όφελος από την ενδεχόμενη κλινική χρήση των συστημάτων βαθμολόγησης για την υποστήριξη αποφάσεων στη ΜΕΘ (π.χ. αποφάσεων παρηγορητικής φροντίδας).

Παρουσιάζονται αναλυτικά οι στατιστικές αυτές μέθοδοι στις παρακάτω ενότητες.

3.6.4. Μέτρα Διαχωριστικής Ικανότητας

Τα μέτρα διαχωριστικής ικανότητας έχουν στόχο να εκτιμήσουν την ικανότητα του προγνωστικού συστήματος βαθμολόγησης να διακρίνει τους ασθενείς που θα πεθάνουν από αυτούς που θα επιζήσουν. Το προγνωστικό σκορ θα πρέπει να αποδίδει υψηλότερη πιθανότητα θανάτου σ' αυτούς που δεν επιβιώνουν συγκριτικά με τους επιζώντες.

Στην περίπτωση που χρησιμοποιείται μια δυαδική μέθοδος πρόγνωσης με την οποία θέτουμε κάποιο όριο αποκοπής ή κατώφλι πρόβλεψης στο προγνωστικό σκορ (π.χ. SAPS > 40), η διαχωριστική ικανότητα μπορεί να αξιολογηθεί με την ευαισθησία και την ειδικότητα της μεθόδου και με άλλα παραγόμενα μέτρα όπως ο δείκτης Youden και το συνολικό ποσοστό ακρίβειας (overall accuracy rate). Προγνωστικά συστήματα που λειτουργούν σε συνεχή αριθμητική κλίμακα αξιολογούνται με τις καμπύλες ROC και με τα εμβαδά των καμπυλών αυτών (Diagnostic accuracy measures).

A. Η Ευαισθησία και η Ειδικότητα

Η ακρίβεια μιας δυαδικής προγνωστικής μεθόδου μπορεί να οριστεί και να αξιολογηθεί αναδρομικά με τη βοήθεια δύο μέτρων:

- 1) Την πιθανότητα που έχει η προγνωστικής μέθοδος να προβλέψει μία έκβαση όταν αυτή πράγματι θα συμβεί, που ονομάζεται Ευαισθησία (Sensitivity) και συμβολίζεται με *Se*.
- 2) Την πιθανότητα που έχει η προγνωστικής μέθοδος να προβλέψει ότι δεν θα συμβεί μια έκβαση που πράγματι δεν πρόκειται να συμβεί, που ονομάζεται Ειδικότητα (Specificity) και συμβολίζεται με *Sp*.

Τα δύο αυτά μέτρα μπορούν να γίνουν πιο κατανοητά με τη βοήθεια του Πίνακα 5 που ακολουθεί:

Πίνακας 5- 2x2 πίνακας συνάφειας της πραγματικής έκβασης των προβλεπόμενων αποτελεσμάτων σε δυαδικό προγνωστικό σύστημα

Πραγματική έκβαση	Αποτέλεσμα πρόβλεψης	
	Θετικό (Y=1)	Αρνητικό (Y=0)
Παρούσα (D=1)	TP	FN
Απούσα (D=0)	FP	TN

όπου **TP = True Positive** (αληθώς θετική πρόβλεψη), **FN = False Negative** (ψευδώς αρνητική πρόβλεψη), **FP = False Positive** (ψευδώς θετική), **TN= True Negative** (αληθώς θετική πρόβλεψη).

Η ευαισθησία, που ονομάζεται και TPR (true positive rate) είναι η πιθανότητα το αποτέλεσμα πρόβλεψης να είναι θετικό δεδομένου ότι η έκβαση είναι παρούσα:

$$Se = TPR = P(Y = 1 | D = 1)$$

Η ειδικότητα, που ονομάζεται και TNR (true negative rate) είναι η πιθανότητα το αποτέλεσμα πρόβλεψης να είναι αρνητικό δεδομένου ότι η έκβαση είναι απύουσα:

$$Sp = TNR = P(Y = 0 \mid D = 0)$$

Ισχύουν οι σχέσεις:

$$FPR = 1 - TNR = 1 - Sp \text{ (false positive rate)}$$

$$FNR = 1 - TPR = 1 - Se \text{ (false negative rate)}$$

Στα προγνωστικά σκορ που εξετάζονται στην παρούσα μελέτη (SAPS II, Apache II, SOFA) το αποτέλεσμα της πρόβλεψης εκφράζεται σε συνεχή αριθμητική κλίμακα. Τα θετικά και αρνητικά αποτελέσματα στην περίπτωση αυτή καθορίζονται από το κατώφλι πρόβλεψης c που θα επιλεγθεί, και το οποίο προφανώς επηρεάζει την ευαισθησία και την ειδικότητα που σε αυτή την περίπτωση εκφράζονται ως εξής:

$$Se = TPR(c) = P(Y \geq c \mid D = 1)$$

$$Sp = TNR(c) = P(Y < c \mid D = 0)$$

Η εκτίμηση των παραπάνω παραμέτρων γίνεται από τις αντίστοιχες δειγματικές αναλογίες στα διαθέσιμα δεδομένα και συνοδεύεται από 95% διάστημα εμπιστοσύνης, το οποίο μπορεί να υπολογιστεί με τη μέθοδο Wald που βασίζεται στην προσέγγιση της κανονικής κατανομής ως εξής:

- 95% διάστημα εμπιστοσύνης για την εκτίμηση της ευαισθησίας:

$$\widehat{Se} \mp 1.96 \sqrt{\frac{\widehat{Se}(1 - \widehat{Se})}{n}}$$

- 95% διάστημα εμπιστοσύνης για την εκτίμηση της ειδικότητας:

$$\widehat{Sp} \mp 1.96 \sqrt{\frac{\widehat{Sp}(1 - \widehat{Sp})}{n}}$$

B. Μέτρα συνολικής ακρίβειας

Η συνολική ακρίβεια (overall accuracy rate) είναι η πιθανότητα σωστής πρόβλεψης (είτε αληθώς θετικής είτε αληθώς αρνητικής) στο σύνολο των ατόμων της μελέτης, δηλαδή το συνολικό ποσοστό των ατόμων που προβλέφθηκε ότι θα αποβιώσουν και όντως απεβίωσαν καθώς και των ατόμων που προβλέφθηκε ότι θα επιζήσουν και όντως επέζησαν. Υπολογίζεται ως ακολούθως:

$$\text{Overall accuracy} = (TP + TN)/(TP + FP + FN + TN)$$

Ένα ακόμη μέτρο συνολικής ακρίβειας της προγνωστικής μεθόδου είναι ο δείκτης του Youden (Youden index), ο οποίος χρησιμεύει για συγκρίσεις μεταξύ πολλών προγνωστικών μεθόδων και ορίζεται ως:

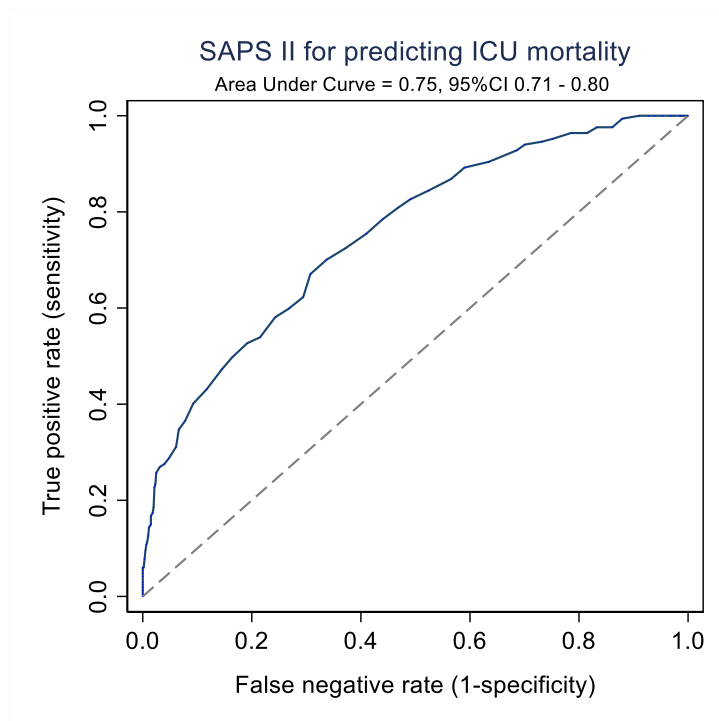
$$YI = Se + Sp - 1 = TPR - FPR$$

Οι τιμές του δείκτη Youden μπορούν να κυμανθούν από 0 έως 1. Τιμή κοντά στο 0 δείχνει ότι η προγνωστική μέθοδος έχει μικρή διαχωριστική ικανότητα, ενώ τιμή κοντά στο 1 υποδεικνύει καλή διαχωριστική ικανότητα. Ο δείκτης Youden είναι επίσης ένα από τα μέτρα που χρησιμοποιούνται για τον προσδιορισμό του βέλτιστου κατωφλίου πρόβλεψης ή απόφασης σε προγνωστικές μεθόδους που παράγουν αποτελέσματα σε συνεχή αριθμητική κλίμακα ώστε να επιτυγχάνεται ο βέλτιστος συνδυασμός ευαισθησίας και ειδικότητας (δίνοντας μέγιστο δείκτη Youden)²¹. Ωστόσο, η επιλογή βέλτιστου κατωφλίου πρόβλεψης με μεγιστοποίηση του δείκτη Youden προϋποθέτει ότι η ευαισθησία και η ειδικότητα θεωρούνται εξίσου σημαντικές.

Γ. Η καμπύλη ROC και το Εμβαδόν της (AUC ή C-index)

Όταν η προγνωστική μέθοδος παράγει αποτελέσματα σε συνεχή αριθμητική κλίμακα, αξιολογούμε την προγνωστική της ακρίβεια με Receiver Operating Characteristic (ROC) Analysis. Η καμπύλη ROC απεικονίζει το σύνολο των δυνατών ζευγών FPR (Se) και TPR ($1 - Sp$) που μπορεί να ληφθούν από όλες τις δυνατές επιλογές για το κατώφλι πρόβλεψης και έχει την ακόλουθη μορφή.

Σχήμα 1- Ενδεικτική καμπύλη ROC από τα αποτελέσματα της μελέτης



Η διαγώνιος (διακεκομμένη γραμμή) στο σχήμα XX μπορεί να ερμηνευτεί ως η καμπύλη ROC που αντιστοιχεί σε ένα προγνωστικό σκορ με μηδενική διαχωριστική ικανότητα όπου η πιθανότητα σωστής πρόβλεψης για οποιοδήποτε κατώφλι c είναι ίση με 0.5, καθώς $FPR(c) = TPR(c)$, δηλαδή εντελώς τυχαία. Στον αντίποδα βρίσκονται τα προγνωστικά συστήματα που έχουν, για κάθε κατώφλι πρόβλεψης, c , τέλεια διαχωριστική ικανότητα, δηλαδή πιθανότητα σωστής πρόβλεψης ίση με 1, και συγκεκριμένα $TPR(c) = 1$ και $FPR(c) = 0$. Η καμπύλη ROC ενός «τέλειου» προγνωστικού σκορ βρίσκεται πάνω στους άξονες y και x και ταυτίζεται με την άνω αριστερή γωνία του σχήματος, δηλαδή αποτελείται από τα ευθύγραμμα τμήματα $(0,0)-(0,1)$ και $(0,1)-(1,1)$. Στην πράξη μια καμπύλη ROC βρίσκεται σε μία ενδιάμεση κατάσταση όπως δείχνει το Σχήμα 1. Όσο πιο πολύ πλησιάζει η καμπύλη ROC την άνω αριστερή γωνία, τόσο καλύτερη είναι η διαχωριστική ικανότητα.

Η καμπύλη ROC είναι ιδιαίτερα εύχρηστη καθώς παρέχει τις τιμές της ευαισθησίας και της ειδικότητας για όλο το φάσμα των πιθανών ορίων απόφασης και δεν εξαρτάται υποκείμενη θνησιμότητα στον υπό μελέτη πληθυσμό ασθενών, επομένως μπορεί να χρησιμοποιηθεί για διαχρονικές και διατοπικές συγκρίσεις, όπως άλλωστε και οι Se και Sp . Η πληροφορία που παρέχει η καμπύλη ROC γίνεται ακόμη πιο εύχρηστη όταν συνοψίζεται με ένα αριθμητικό μέτρο. Τέτοιος δείκτης είναι το εμβαδόν κάτω από

την καμπύλη (AUC) που είναι ιδιαίτερα χρήσιμο όταν πρόκειται για τη σύγκριση πολλών προγνωστικών σκορς. Επειδή το εμβαδόν κάτω από τη διαγώνιο είναι 0.5 και το συνολικό εμβαδόν του τετραγώνου είναι 1, ισχύει

$$0.5 \leq AUC \leq 1$$

Ένα προγνωστικό σκορ με $AUC = 0.5$ είναι πρακτικά άχρηστο αφού η πρόβλεψη της έκβασης των ασθενών ισοδυναμεί με τη ρίψη ενός νομίσματος (είναι εντελώς τυχαία). Αντίθετα, ένα προγνωστικό σκορ με τέλεια διαχωριστική ικανότητα θεωρητικά θα έχει $AUC = 1$.

Επειδή εκτιμάται το εμβαδόν AUC από τα διαθέσιμα δειγματικά δεδομένα, μας ενδιαφέρει η στατιστική σημαντικότητα της διαφοράς από το 0.5 ($H_0: AUC = 0.5$). Αυτό μπορούμε να το διαπιστώσουμε εύκολα υπολογίζοντας ένα 95% διάστημα εμπιστοσύνης για το AUC. Στο σχήμα XX, για παράδειγμα, έχουμε $\widehat{AUC} = 0.75$ με 95% διάστημα εμπιστοσύνης 0.71 – 0.80, οπότε η AUC είναι στατιστικά σημαντική ($p < 0.05$).

Σημειώνουμε ότι για δυαδικές εκβάσεις, το AUC της ROC ταυτίζεται με το στατιστικό δείκτη concordance index (c-index) και συχνά στη βιβλιογραφία αναφέρεται με τη δεύτερη ονομασία.

Επίσης, στην περίπτωση δυαδικού προγνωστικού συστήματος (με επιλογή κάποιου ορίου αποκοπής), ισχύει:

$$AUC = \frac{Se + Sp}{2}$$

3.6.5. Μέτρα Βαθμονόμησης

Βαθμονόμηση είναι η αντιστοιχία που υπάρχει μεταξύ των προβλέψεων του προγνωστικού συστήματος και των παρατηρούμενων αποτελεσμάτων, δηλαδή το βαθμό συμφωνίας ανάμεσα στα παρατηρούμενα ποσοστά θνησιμότητας (ή άλλης έκβασης) και στις προβλεπόμενες πιθανότητες θανάτου σύμφωνα με τα προγνωστικά συστήματα βαθμολόγησης.

Κατά την αξιολόγηση ενός προβλεπτικού σκορ, μας ενδιαφέρει να υπάρχει ακρίβεια στη συνολική βαθμονόμηση (calibration-in-the-large). Για παράδειγμα, το προγνωστικό σκορ προβλέπει συνολική (μέση) θνησιμότητα 70% στους ασθενείς της μελέτης, η βαθμονόμηση είναι σωστή αν η παρατηρούμενη θνησιμότητα είναι 70%. Ωστόσο, ακόμη και όταν η συνολική βαθμονόμηση είναι

ακριβής, ενδέχεται να υπάρχουν σοβαρές αποκλίσεις σε συγκεκριμένες ομάδες ατόμων (π.χ. σε ασθενείς υψηλού κινδύνου). Επομένως, είναι σημαντικό να αξιολογείται η βαθμονόμηση σε όλο το εύρος της θνησιμότητας και όχι απλώς κατά μέσο όρο²².

Συχνά χρησιμοποιούμενες στατιστικές μέθοδοι για την αξιολόγηση της βαθμονόμησης είναι ο τυποποιημένος δείκτης θνησιμότητας (Standardized Mortality Ratio - SMR), ο στατιστικός έλεγχος καλής προσαρμογής των Hosmer-Lemeshow, και τα διαγράμματα βαθμονόμησης (calibration plots).

A. Υπολογισμός των Προβλεπόμενων Πιθανοτήτων Θανάτου

Ο υπολογισμός της προβλεπόμενης πιθανότητας θανάτου για τα προγνωστικά σκορ της παρούσας μελέτης (SAPS II, Apache II, και SOFA) έγινε με τη βοήθεια του μοντέλου λογιστικής παλινδρόμησης. Το μοντέλο αυτό έχει τη μορφή:

$$\log Odds = a + b X + \varepsilon$$

Στην εξίσωση αυτή, *Odds* είναι το πηλίκο συμπληρωματικών πιθανοτήτων θανάτου, δηλαδή

$$Odds = \frac{P(D = 1)}{1 - P(D = 1)} = \frac{P(D = 1)}{P(D = 0)}$$

όπου $P(D = 1)$ είναι η προβλεπόμενη πιθανότητα θανάτου και $P(D = 0)$ η προβλεπόμενη πιθανότητα επιβίωσης.

Η μεταβλητή X είναι το προγνωστικό σκορ (ένα από τα SAPS II, Apache II, ή SOFA). Ο όρος ε παριστάνει το τυχαίο σφάλμα, δηλαδή τυχαίες αποκλίσεις από το μοντέλο που θεωρείται ότι κατά μέσο όρο είναι μηδενικές. Τέλος, \log είναι ο φυσικός λογάριθμος (\ln) με βάση τον άρρητο αριθμό e .

Ο συντελεστής b ποσοτικοποιεί τη συσχέτιση της έκβασης με το προγνωστικό σκορ X και ειδικότερα ισχύει $e^b = Odds Ratio (OR)$.

Από την προηγούμενη εξίσωση:

$$\log \frac{P(D = 1)}{1 - P(D = 1)} = a + b X + \varepsilon$$

Αντι-λογαριθμίζοντας (παίρνοντας εκθετικά e) και στα δύο μέλη, έχουμε:

$$\frac{P(D = 1)}{1 - P(D = 1)} = e^{a + bX + \varepsilon}$$

και επιλύοντας την εξίσωση ως προς την προβλεπόμενη πιθανότητα $P(D = 1)$ λαμβάνεται:

$$P(D = 1) = \frac{e^{a + bX + \varepsilon}}{1 + e^{a + bX + \varepsilon}}$$

Οι συντελεστές a και b υπολογίζονται από τα διαθέσιμα δεδομένα με τη μέθοδο της μέγιστης πιθανοφάνειας (maximum likelihood estimation). Αν \hat{a} και \hat{b} είναι οι δειγματικές εκτιμήσεις, τότε η προβλεπόμενη πιθανότητα θανάτου εκτιμάται από το δείγμα της μελέτης ως:

$$P(\widehat{D} = 1) = \frac{e^{\hat{a} + \hat{b}X}}{1 + e^{\hat{a} + \hat{b}X}}$$

Στην παραπάνω εκτίμηση έχει παραληφθεί ο όρος σφάλματος ε αφού αναμένεται το σφάλμα αυτό να είναι μηδενικό. Ωστόσο, η αβεβαιότητα εξαιτίας τυχαίας δειγματοληπτικής διακύμανσης εκφράζεται μέσω του διαστήματος εμπιστοσύνης για την $P(D = 1)$.

B. Ο Τυποποιημένος Δείκτης Θνησιμότητας (SMR)

Πρόκειται για το πηλίκο των παρατηρούμενων προς τους αναμενόμενους θανάτους του δείγματος. Ο αναμενόμενος αριθμός θανάτων προκύπτει αθροίζοντας τις ατομικές πιθανότητες θανάτου των ασθενών όπως αυτές υπολογίζονται από τα προγνωστικά σκορς. Τιμή του SMR ίση με 1 υποδηλώνει τέλεια συνολική βαθμονόμηση του συστήματος, καθώς δείχνει απόλυτη συμφωνία μεταξύ των συνολικών παρατηρούμενων και των αναμενόμενων θανάτων. Τιμή του SMR μεγαλύτερη της μονάδας υποδηλώνει υποεκτίμηση των θανάτων καθώς οι παρατηρούμενοι θάνατοι είναι περισσότεροι από τους αναμενόμενους, ενώ τιμή μικρότερη της μονάδας υποδηλώνει υπερεκτίμηση των θανάτων καθώς οι παρατηρούμενοι θάνατοι είναι λιγότεροι από τους αναμενόμενους.

Γ. Ο Έλεγχος Καλής Προσαρμογής των Hosmer και Lemeshow

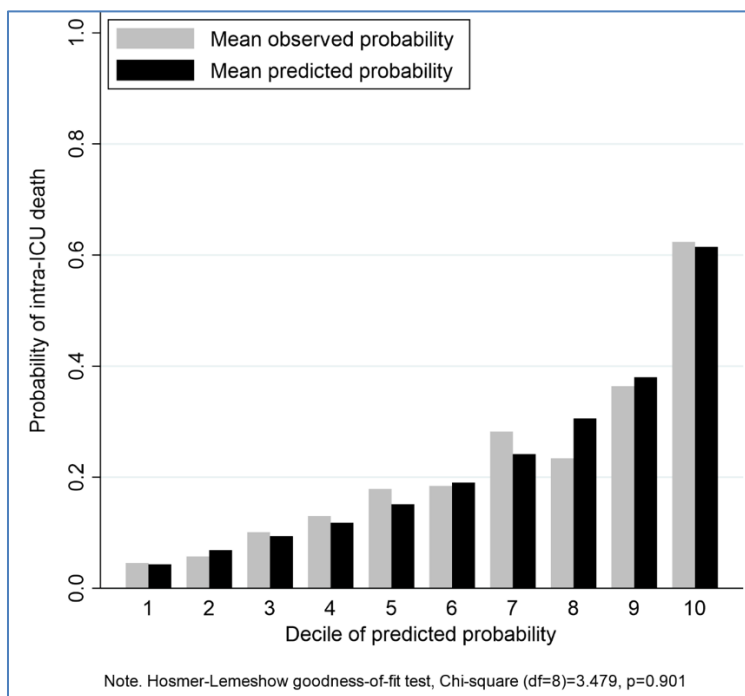
Ο στατιστικός έλεγχος υπολογίζει τις αναμενόμενους αριθμούς θανάτων των ασθενών, κατηγοριοποιώντας τις σε ομάδες που καθορίζονται από τις προβλεπόμενες από το προγνωστικό σκορ πιθανότητες θανάτου. Η πιο συχνή ομαδοποίηση βασίζεται στα δεκατημόρια των προβλεπόμενων πιθανοτήτων (δηλαδή σε 10 ομάδες, ανά 10%). Η σύγκριση μεταξύ των παρατηρούμενων και προβλεπόμενων συχνοτήτων των ασθενών που πεθαίνουν και επιβιώνουν σε κάθε δεκατημόριο αποτελεί τη βάση του ελέγχου για την εκτίμηση της προσαρμογής του προγνωστικού μοντέλου στα δεδομένα της μελέτης. Το αποτέλεσμα του ελέγχου συγκρίνεται με τη χ^2 κατανομή με 8 βαθμούς ελευθερίας (γενικά οι β.ε. είναι $k - 2$, όπου k είναι ο αριθμός των ομάδων) ώστε να υπολογιστεί η τιμή πιθανότητας (p -value) του ελέγχου. Η μηδενική υπόθεση στον έλεγχο καλής προσαρμογής είναι ότι δεν υπάρχει διαφορά ανάμεσα στα παρατηρούμενα ποσοστά θανάτων και τις αντίστοιχες προβλεπόμενες πιθανότητες (H_0 : observed risks = predicted risks) και οι όποιες διαφορές παρατηρήθηκαν στην μελέτη ήταν προϊόντα τυχαίας δειγματοληπτικής διακύμανσης. Μικρή αντιστοιχία μεταξύ προβλεπόμενων και παρατηρούμενων τιμών αντιστοιχεί σε μεγάλη τιμή του στατιστικού κριτηρίου χ^2 του ελέγχου και μικρή p -τιμή. Θεωρούμε ότι υπάρχει καλή βαθμονόμηση όταν δεν απορρίπτεται η H_0 σε επίπεδο σημαντικότητας $\alpha=0.05$.

Δ. Ραβδόγραμμα Βαθμονόμησης

Ο έλεγχος των Hosmer – Lemeshow μπορεί να απεικονιστεί γραφικά με τη μορφή της ραβδογράμματος βαθμονόμησης, το οποίο συγκρίνει τις παρατηρούμενες και τις προβλεπόμενες πιθανότητες θανάτου ανά δεκατημόριο προβλεπόμενων πιθανοτήτων. Ένα παράδειγμα δίνεται στο σχήμα Σχήμα 2.

Στο διάγραμμα αυτό είναι σημαντικό να ελεγχθεί ότι δεν υπάρχει κάποιο συστηματικό μοτίβο απόκλισης των συγκρινόμενων πιθανοτήτων, όπως για παράδειγμα να υποεκτιμούνται οι πιθανότητες στους ασθενείς χαμηλού κινδύνου και να υπερεκτιμούνται για τους ασθενείς υψηλού κινδύνου (ή αντίστροφα).

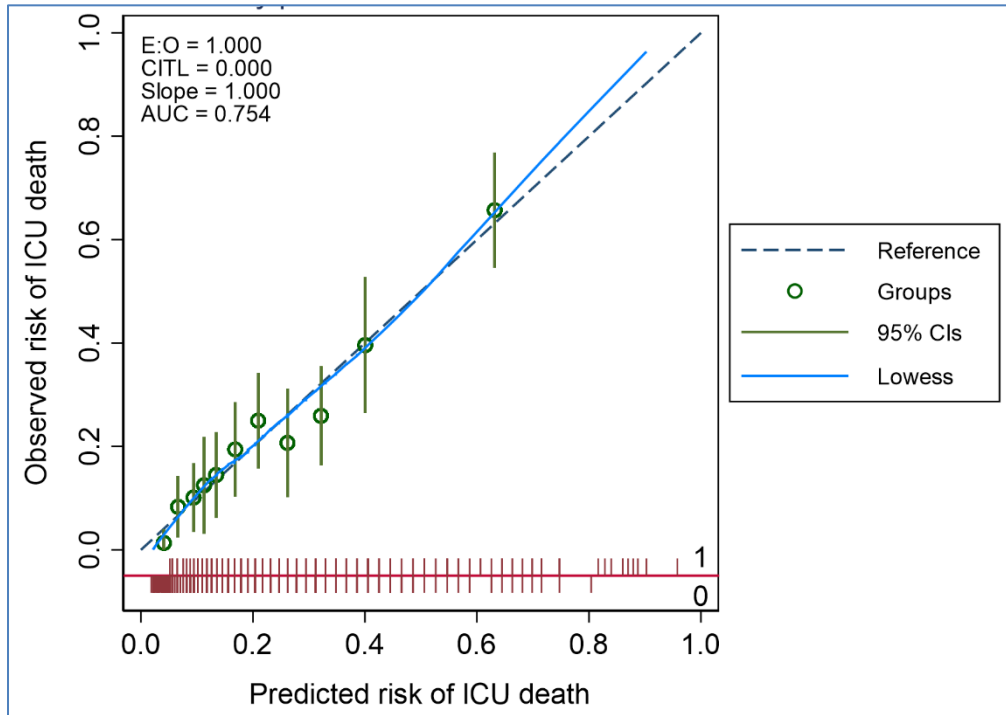
Σχήμα 2- Ενδεικτικό ραβδόγραμμα βαθμονόμησης από τα αποτελέσματα της μελέτης



E. Εξομαλυμένο Διάγραμμα Βαθμονόμησης

Πρόκειται για ένα διάγραμμα διασποράς, όπου ο ένας άξονας (συνήθως ο y) παριστάνει τις παρατηρούμενες πιθανότητες θανάτου και ο άλλος άξονας (συνήθως ο x) τις προβλεπόμενες πιθανότητες από το προγνωστικό σκορ για κάθε ασθενή της μελέτης. Εφόσον οι δύο άξονες παριστάνουν πιθανότητες από 0 έως 1, το διάγραμμα περικλείεται από ένα τετράγωνο. Η διαγώνιος του τετραγώνου (ευθεία με κλίση ίση με 1) αντιπροσωπεύει την τέλεια βαθμονόμηση, όπου κάθε προβλεπόμενη πιθανότητα είναι ακριβώς ίση με την παρατηρούμενη πιθανότητα. Επομένως, είναι ενδιαφέρον ο βαθμός στον οποία τα σημεία του διαγράμματος διασποράς έχουν κάποια τάση να αποκλίνουν από την διαγώνιο. Αυτό μπορεί να παρατηρηθεί προσαρμόζοντας την ευθεία της παλινδρόμησης ελαχίστων τετραγώνων αλλά συχνά είναι περισσότερο κατατοπιστικό να προσαρμόσουμε μια εξομαλυμένη γραμμή lowess²². Στο διάγραμμα αυτό γίνεται ταυτόχρονα να απεικονιστούν τα σημεία (πιθανότητες) που αντιστοιχούν στα δεκατημόρια με τα αντίστοιχα 95% διαστήματα εμπιστοσύνης. Ένα παράδειγμα δίνεται στο Σχήμα 3.

Σχήμα 3- Ενδεικτικό εξομαλυμένο διάγραμμα βαθμονόμησης από τα αποτελέσματα της μελέτης



3.6.6. Καθαρό Όφελος και Καμπύλη Απόφασης

Ένα προγνωστικά σκορ με καλύτερη διαχωριστική ικανότητα και βαθμονόμηση θα πρέπει θεωρητικά να είναι αποτελεί καλύτερο “βοηθητικό οδηγό” στην κλινική διαχείριση των ασθενών. Ωστόσο τα στατιστικά μέτρα διαχωρισμού και βαθμονόμησης υπολείπονται όταν χρειάζεται να αξιολογηθεί αν ένα προγνωστικό σκορ θα μπορούσε να βελτιώσει τη λήψη κλινικών αποφάσεων στην πράξη. Τα προαναφερθέντα μέτρα δεν μπορούν να πληροφορήσουν εάν θα είναι ωφέλιμο να χρησιμοποιηθεί το προγνωστικό σκορ για τη λήψη κλινικών αποφάσεων ή ποιο από δύο ή περισσότερα σκορς θα οδηγήσει σε καλύτερες αποφάσεις, ειδικά εάν το ένα προγνωστικό σκορ έχει καλύτερη διαχωριστική ικανότητα και το άλλο καλύτερη βαθμονόμηση²³.

Για να ξεπεραστεί αυτός ο περιορισμός, προτάθηκε από τους Vickers και Elkin το 2006 η χρήση του καθαρού οφέλους (NB – net benefit) ως μέτρου ικανού να συνοψίσει την απόδοση ενός προγνωστικού μοντέλου για την υποστήριξη στη λήψη αποφάσεων¹⁸. Οι εκτιμήσεις του καθαρού οφέλους δίνονται για ένα εύρος δυνατών πιθανοτήτων κατωφλίου απόφασης σε γραφικές παραστάσεις και η μέθοδος είναι γνωστή ως Ανάλυση Καμπύλης Απόφασης (DCA – Decision Curve Analysis).

A. Η πιθανότητα κατωφλίου απόφασης (p_t)

Για να χρησιμοποιηθεί ένα προγνωστικό σκορ στην κλινική πράξη, ορίζεται μια πιθανότητα κατωφλίου (threshold probability ή risk threshold) πάνω από την οποία θα αποφασιστεί η κλινική “διαχείριση ή παρέμβαση” (“treatment”)²³.

Πρέπει να σημειωθεί ότι ο όρος “παρέμβαση” (ή “διαχείριση”) αναφέρεται σε ένα ευρύ φάσμα δυνατών δράσεων, όπως πρόσθετη διαγνωστική διερεύνηση (π.χ. βιοψία), παραπομπή σε εξειδικευμένη φροντίδα, χειρουργική επέμβαση, ή ιατρική θεραπεία. Ιδιαίτερο ενδιαφέρον σε σχέση με το υλικό της παρούσας μελέτης έχουν οι αποφάσεις στη ΜΕΘ που λαμβάνονται σε καταστάσεις τέλους ζωής, όπου δίδεται προτεραιότητα στη βελτίωση της ποιότητας ζωής του ασθενούς και λιγότερο στην παράταση της. Ο στόχος αυτός μπορεί να απαιτεί την εφαρμογή υποστηρικτικής αγωγής ή την αποκλιμάκωση της θεραπευτικής αγωγής, εάν αυτή δεν ωφελεί ή αν περιορίζονται οι παρενέργειες. Έχει αναδειχθεί ότι ο καθορισμός εκ των προτέρων των στόχων της θεραπείας του βαρέως πάσχοντα εντός 72 ωρών από την εισαγωγή του στη ΜΕΘ, προωθεί την επικοινωνία και την κατανόηση μεταξύ της οικογένειας του ασθενούς και της θεραπευτικής ομάδας, μειώνοντας το στρες και την κατάθλιψη των επιζώντων συγγενών και τελικά διευκολύνει τη διαδικασία λήψης αποφάσεων σε καταστάσεις τέλους της ζωής²⁴⁻²⁶. Βεβαίως, για να είναι δυνατή η ύπαρξη ρεαλιστικών στόχων θεραπείας στη ΜΕΘ, θα πρέπει το αποτέλεσμα της εντατικής θεραπείας σε βαρέως πάσχοντες ασθενείς να μπορεί να προβλεφθεί με ακρίβεια.

Ανάλογα με την πιθανότητα κατωφλίου που θα χρησιμοποιηθεί για τη λήψη απόφασης, ενδέχεται να ληφθούν λανθασμένες αποφάσεις “παρέμβασης”, οι οποίες μπορεί να είναι ψευδώς θετικές (TP) και ψευδώς αρνητικές (FN). Για παράδειγμα, μπορεί να υπάρξει βλάβη από άσκοπη θεραπεία (ψευδώς θετικά) σε σύγκριση με το όφελος της θεραπείας σε εκείνους που πραγματικά τη χρειάζονται (αληθώς θετικά). Πολλοί ερευνητές επιλέγουν την πιθανότητα κατωφλίου απόφασης με τρόπο ώστε να

μεγιστοποιείται το άθροισμα $TP + TN$ των αληθώς θετικών και των αληθώς αρνητικών αποφάσεων, δηλαδή το άθροισμα ευαισθησίας και ειδικότητας (ο δείκτης Yuden). Ωστόσο, αυτό προϋποθέτει ότι μια αληθώς θετική απόφαση θεωρείται εξίσου σημαντική με μια αληθώς αρνητική απόφαση και αυτό μπορεί να μην συμβαίνει στην πράξη (π.χ. μπορεί να είναι πιο σημαντικό να βρεθεί ένας επιθετικός καρκίνος παρά να αποφευχθεί μια περιττή βιοψία).

Σύμφωνα με τη θεωρία αποφάσεων, η πιθανότητα κατωφλίου απόφασης p_t αντανakλά το σημείο στο οποίο αδιαφορούμε για την «παρέμβαση», δηλαδή εκεί όπου το αναμενόμενο όφελος από την «παρέμβαση» είναι ίσο με το αναμενόμενο όφελος από την αποφυγή της «παρέμβασης»^{23,27}. Εάν το προγνωστικό σκορ δίνει πιθανότητα έκβασης που είναι μικρότερη από το κατώφλι απόφασης, $P(\widehat{D} = 1) < p_t$, τότε ο ασθενής θεωρείται «αρνητικός» και η ενδεδειγμένη απόφαση είναι να αποφευχθεί η παρέμβαση. Διαφορετικά, εάν $P(\widehat{D} = 1) > p_t$ τότε ο ασθενής θεωρείται «θετικός» και η ενδεδειγμένη απόφαση είναι να γίνει η παρέμβαση. Όπως τονίσαμε και παραπάνω, το κατώφλι απόφασης πρέπει να αποφασιστεί σταθμίζοντας ανάμεσα στο όφελος από μια ορθή «παρέμβαση» (σε ασθενείς που την χρειάζονται) και το κόστος από μια περιττή «παρέμβαση» (σε ασθενείς που δεν την χρειάζονται). Για παράδειγμα, ένας ασθενής (ή η οικογένειά του) μπορεί να επιλέξει να λάβει την «παρέμβαση» αν πηλίκο (odds) οφέλους προς κόστος είναι 4, δηλαδή αν το όφελος από μια χρήσιμη παρέμβαση είναι 4-πλάσιο από το κόστος μιας περιττής παρέμβασης. Αυτό σημαίνει ότι

$$\frac{p_t}{1 - p_t} = 4 \rightarrow p_t = 0.80$$

Διαφορετικοί ασθενείς (ή οι οικογένειες του ή οι γιατροί τους) μπορούν να έχουν διαφορετικές προτιμήσεις, επομένως πρέπει κανείς να θεωρεί ένα εύρος δυνατών τιμών για την πιθανότητα κατωφλίου απόφασης όταν αξιολογεί ένα ή περισσότερα προγνωστικά σκορς.

B. Το καθαρό όφελος (NB)

Οι Vickers και Elkin το 2006 όρισαν το καθαρό όφελος από τη χρήση ενός προγνωστικού συστήματος με τρόπο παρόμοιο του καθαρού κέρδους σε μια επιχείρηση: έσοδα μείον δαπάνες¹⁸. Κατ' αναλογία, τα έσοδα προέρχονται από τις αληθώς θετικές αποφάσεις (χρήσιμες παρεμβάσεις) και οι δαπάνες προκαλούνται από τις ψευδώς θετικές αποφάσεις (περιττές παρεμβάσεις). Αλλά, αυτά τα «έσοδα» και

οι “δαπάνες” είναι σε διαφορετική κλίμακα – έχουν διαφορετική βαρύτητα – που αντανακλά το πηλίκο του οφέλους από χρήσιμες παρεμβάσεις και του κόστους από περιττές παρεμβάσεις $\frac{p_t}{1-p_t}$. Κατά συνέπεια, το καθαρό όφελος υπολογίζεται ως εξής:

$$NB = \frac{TP(p_t)}{n} - \frac{p_t}{1-p_t} \cdot \frac{FP(p_t)}{n}$$

όπου

n είναι ο συνολικός αριθμός των ασθενών στη μελέτη,

$TP(p_t)$ είναι ο αριθμός των ασθενών με αληθώς θετικές αποφάσεις σε κατώφλι p_t , και

$FP(p_t)$ είναι ο αριθμός των ασθενών με ψευδώς θετικές αποφάσεις σε κατώφλι p_t .

Πολλαπλασιαζόμενη επί 100, η τιμή του NB μπορεί να ερμηνευτεί ως ο αριθμός των χρήσιμων παρεμβάσεων (που αναγνωρίστηκαν εξαιτίας της χρήσης του προγνωστικού σκορ) σε 100 ασθενείς, οι οποίοι θα είχαν (χωρίς τη χρήση του προγνωστικού σκορ) αφεθεί χωρίς παρέμβαση.

Αν, για παράδειγμα, η έκβαση είναι ο θάνατος στη ΜΕΘ, η παρέμβαση είναι η έναρξη παρηγορητικής φροντίδας και υπολογιστεί ότι η χρήση ενός προγνωστικού σκορ επιφέρει $NB=0.06$, αυτό σημαίνει ότι 6 ανά 100 ασθενείς που πρόκειται να πεθάνουν στη ΜΕΘ θα τοποθετηθούν σε παρηγορητική φροντίδα, χωρίς να υπάρξουν περιττές / αχρείαστες τοποθετήσεις.

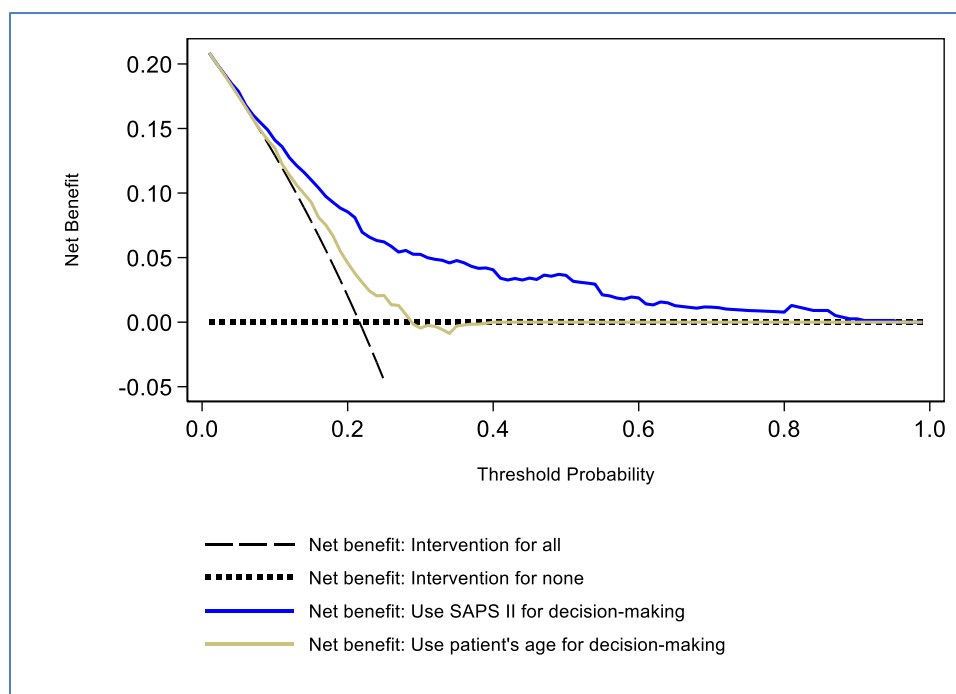
Γ. Η Καμπύλη Απόφασης

Η πιθανότητα κατωφλίου απόφασης (p_t) στις κλινικές εφαρμογές δεν είναι συχνά ορισμένη με ακρίβεια ή δεν υπάρχει κάποια καθολικά αποδεκτή τιμή. Η σχετική βαρύτητα οφέλους και κόστους μπορεί να μην είναι γνωστή γιατί δεν υπάρχουν διαθέσιμα επιστημονικά δεδομένα ή/και εξαιτίας διαφορετικών εκτιμήσεων των ιατρών ή προτιμήσεων των ασθενών. Επομένως, όπως τονίσαμε και προηγουμένα, είναι απαραίτητο να θεωρούμε ένα εύρος δυνατών τιμών για την πιθανότητα κατωφλίου απόφασης όταν αξιολογούμε ένα ή περισσότερα προγνωστικά σκορς. Αυτό οδηγεί στον

υπολογισμό ενός αντίστοιχου εύρους τιμών για το NB (μία τιμή NB για κάθε κατώφλι p_t), οι οποίες παριστάνονται σε μια καμπύλη απόφασης, δηλαδή ένα γράφημα των σημείων (p_t , NB).

Για να ερμηνεύσουμε σωστά το NB, εισάγουμε δύο ακραίες στρατηγικές στις οποίες η διαχείριση των ασθενών γίνεται χωρίς τη χρήση του προγνωστικού μοντέλου¹⁸. Η μία στρατηγική είναι να αποφευχθεί η παρέμβαση για όλους τους ασθενείς (treat none), και η άλλη στρατηγική είναι να υλοποιηθεί η παρέμβαση για όλους τους ασθενείς (treat all). Στην πρώτη περίπτωση έχουμε NB = 0 γιατί δεν λαμβάνεται καμία δράση, ενώ το NB για τη δεύτερη περίπτωση μεταβάλλεται αναλόγως του κατωφλίου πιθανότητας απόφασης p_t . Ένα παράδειγμα απεικονίζεται στο Σχήμα 4.

Σχήμα 4- Ενδεικτική καμπύλη απόφασης τα αποτελέσματα της μελέτης



3.7. Έγκριση από Επιτροπή Βιοηθικής

Η μελέτη έλαβε έγκριση από την αρμόδια επιστημονική επιτροπή του Πανεπιστημιακού Γενικού Νοσοκομείου, καθώς και από την 7^η ΔΥΠΕ Κρήτης (Α.Π. 6377).

4. Αποτελέσματα

4.1. Χαρακτηριστικά πληθυσμού μελέτης

Κατά τη διετία της μελέτης, εισήχθησαν στη ΜΕΘ ενηλίκων του ΠαΓΝΗ 778 ασθενείς, με διάμεση ηλικία 68 έτη (εύρος 17 - 98, εκ των οποίων οι 498 (64.0%) ήταν άνδρες. Η θνησιμότητα εντός της ΜΕΘ ανήλθε σε 170 άτομα, ήτοι 21.9% (95%ΔΕ 19.0 - 24.9 %), σε μέσο χρόνο 11.1±13.6 ημέρες. Η συνολική ενδονοσοκομειακή θνησιμότητα ανήλθε σε 270 άτομα, ήτοι 34.7% (95%ΔΕ 31.4 - 38.2 %).

Ο Πίνακας 6 συνοψίζει κλινικά χαρακτηριστικά των ασθενών, συμπεριλαμβανομένης της κύριας διάγνωσης εισαγωγής, της ταξινόμησης και της προέλευσης των περιστατικών. Οι μέσες τιμές των προγνωστικών κλιμάκων SAPS II, APACHE II και SOFA κατά τις πρώτες 24 ώρες εισαγωγής στη ΜΕΘ ήταν αντίστοιχα: 40.9±13.0, 23.6±7.4 και 9.9±3.5. Τα δεδομένα ήταν ελλιπή για τον υπολογισμό των δεικτών αυτών σε 17 (2.2%) ασθενείς για το SAPS II, 20 (2.6%) ασθενείς για το APACHE II και 20 (2.6%) ασθενείς για το SOFA.

Τα ιστογράμματα των κατανομών των προγνωστικών κλιμάκων καθώς και της ηλικίας και της διάρκειας νοσηλείας των ασθενών δίνονται στο συμπληρωματικό **Σχήμα S1**, στο παράρτημα.

Στον **Πίνακα 7** συγκρίνονται τα κλινικά χαρακτηριστικά και οι προγνωστικές κλίμακες (SAPS II, APACHE II, SOFA) ανάμεσα σε ασθενείς που επιβίωσαν και ασθενείς που απεβίωσαν στη ΜΕΘ. Παρόμοια ανάλυση γίνεται στον **Πίνακα 8** σε σχέση με τη συνολική νοσοκομειακή θνησιμότητα. Το συμπληρωματικό **Σχήμα S2** δείχνει τα συγκριτικά ιστογράμματα για τις συνεχείς μεταβλητές της μελέτης.

Όπως ήταν αναμενόμενο, οι ασθενείς που απεβίωσαν είχαν κατά μέσο όρο στατιστικά σημαντικά μεγαλύτερη ηλικία, όσο και σκορς SAPS II, APACHE II και SOFA κατά την εισαγωγή τους στη ΜΕΘ. Επίσης, η πιθανότητα θανάτου στη ΜΕΘ ήταν αυξημένη για ασθενείς που εισήχθησαν με νοσήματα του κυκλοφορικού και του αναπνευστικού συστήματος, για εκείνους που εισήχθησαν με λοιμώξεις, και για τα παθολογικά περιστατικά γενικότερα. Αντιθέτως, υψηλές πιθανότητες επιβίωσης είχαν οι ασθενείς κάτω των 55 ετών, οι τραυματίες και οι χειρουργικοί ασθενείς.

Πίνακας 6. Κλινικά χαρακτηριστικά των ασθενών (n = 778) κατά την εισαγωγή στη ΜΕΘ, έκβαση στη ΜΕΘ και έκβαση στο νοσοκομείο

Χαρακτηριστικό	Στατιστικές σύνοψης¹
Φύλο Άρρεν, n(%)	498 (64.0%)
Ηλικία (έτη), MT ± TA [ΔΜ (IQR)]	64.4 ± 17.3 [68.0 (55.0-77.0)]
Νοσηλεία πριν τη ΜΕΘ (ημέρες), MT ± TA [ΔΜ (IQR)]	6.6 ± 24.1 [1.0 (0.0-5.0)]
Κυρία διάγνωση εισαγωγής στη ΜΕΘ, n(%)	
Νόσημα κυκλοφορικού συστήματος	181 (23.3%)
Νόσημα αναπνευστικού συστήματος	134 (17.2%)
Νεόπλασμα	102 (13.1%)
Τραυματισμός ή εξωτερική αιτία	87 (11.2%)
Νόσημα πεπτικού συστήματος	68 (8.7%)
Λοιμώδες νόσημα	33 (4.2%)
Νόσος νευρικού συστήματος	28 (3.6%)
Νόσημα ουρογεννητικού συστήματος	15 (1.9%)
Άλλη νόσος	130 (16.7%)
Ταξινόμηση περιστατικού, n(%)	
Επείγον χειρουργείο	170 (21.9%)
Προγραμματισμένο Χειρουργείο	94 (12.1%)
Παθολογικό περιστατικό	425 (54.6%)
Τραύμα	89 (11.4%)
Προέλευση περιστατικού, n(%)	
Κλινική/ Άλλη ΜΕΘ	398 (51.2%)
ΤΕΠ/Διακομιδή	279 (35.9%)
Χειρουργείο	101 (13.0%)
SAPS II, MT ± TA [ΔΜ (IQR)]	40.9 ± 13.0 [40.0 (32.0-49.0)]
APACHE II, MT ± TA [ΔΜ (IQR)]	23.6 ± 7.4 [23.0 (18.0-28.0)]
SOFA, MT ± TA [ΔΜ (IQR)]	9.9 ± 3.5 [10.0 (8.0-12.0)]
Παραμονή στη ΜΕΘ (ημέρες), MT ± TA [ΔΜ (IQR)]	11.1 ± 13.6 [6.0 (3.0-15.0)]
Θνησιμότητα στη ΜΕΘ, n(%) ²	170 (21.9%)
Επανεισαγωγή στη ΜΕΘ, n(%)	63 (8.1%)
Συνολική ενδονοσοκομειακή θνησιμότητα, n(%)	270 (34.7%)
Τελική έκβαση στο νοσοκομείο, n(%)	
Θάνατος	270 (34.7%)
Επιδείνωση	4 (0.5%)
Στασιμότητα	23 (3.0%)
Βελτίωση	465 (59.8%)
Ίαση	16 (2.1%)
Συνολικής νοσηλεία (ημέρες), MT ± TA [ΔΜ (IQR)]	33.3 ± 41.3 [21.5 (12.0-38.0)]

Σημειώσεις:

¹ MT = μέση τιμή, TA = τυπική απόκλιση, ΔΜ = διάμεσος, IQR = ενδοτεταρτημοριακό εύρος

² Η θνησιμότητα στη ΜΕΘ περιλαμβάνει θανάτους που συνέβησαν μετά από επανεισαγωγή στη ΜΕΘ.

Πίνακας 7. Σύγκριση κλινικών χαρακτηριστικών και προγνωστικών δεικτών (SAPS II, APACHE II, SOFA) των ασθενών που επιβίωσαν και εκείνων που απεβίωσαν στη ΜΕΘ

	Επιβίωσαν N=608	Απεβίωσαν N=170	p-value
Φύλο, n(%)			0.61
Θήλυ	216 (77.1)	64 (22.9)	
Άρρεν	392 (78.7)	106 (21.3)	
Ηλικία (έτη), ΔΜ (IQR)	66.5 (52.0-76.0)	73.0 (62.0-80.0)	<0.001
Ηλικιακή ομάδα ανά τεταρτημόριο (έτη) , n(%)			<0.001
17 - 55	178 (89.0)	22 (11.0)	
55 - 68	160 (77.7)	46 (22.3)	
68 - 77	144 (76.2)	45 (23.8)	
77 - 98	126 (68.9)	57 (31.1)	
Νοσηλεία πριν τη ΜΕΘ (ημέρες), MT ± TA	1.0 (0.0-4.0)	1.0 (0.0-8.0)	0.031
Κυρία διάγνωση εισαγωγής στη ΜΕΘ, n(%)			<0.001
Νόσημα κυκλοφορικού συστήματος	123 (68.0)	58 (32.0)	
Νόσημα αναπνευστικού συστήματος	107 (79.9)	27 (20.1)	
Νεόπλασμα	86 (84.3)	16 (15.7)	
Τραυματισμός ή εξωτερική αιτία	79 (90.8)	8 (9.2)	
Νόσημα πεπτικού συστήματος	56 (82.4)	12 (17.6)	
Λοιμώδες νόσημα	22 (66.7)	11 (33.3)	
Νόσος νευρικού συστήματος	23 (82.1)	5 (17.9)	
Νόσημα ουρογεννητικού συστήματος	14 (93.3)	1 (6.7)	
Άλλη νόσος	98 (75.4)	32 (24.6)	
Ταξινόμηση περιστατικού, n(%)			<0.001
Επείγον χειρουργείο	142 (83.5)	28 (16.5)	
Προγραμματισμένο Χειρουργείο	85 (90.4)	9 (9.6)	
Παθολογικό περιστατικό	301 (70.8)	124 (29.2)	
Τραύμα	80 (89.9)	9 (10.1)	
Πρόελευση περιστατικού, n(%)			0.033
Κλινική/ Άλλη ΜΕΘ	304 (76.4)	94 (23.6)	
ΤΕΠ/Διακομιδή	215 (77.1)	64 (22.9)	
Χειρουργείο	89 (88.1)	12 (11.9)	
SAPS II, MT ± TA	38.2 ± 11.2	50.8 ± 14.2	<0.001
APACHE II, MT ± TA	22.4 ± 6.9	28.0 ± 7.4	<0.001
SOFA, MT ± TA	9.4 ± 3.3	12.0 ± 3.2	<0.001
Παραμονή στη ΜΕΘ (ημέρες), ΔΜ (IQR)	6.0 (3.0-14.0)	8.5 (3.0-17.0)	0.024

Σημειώσεις:

¹ MT = μέση τιμή, TA = τυπική απόκλιση, ΔΜ = διάμεσος, IQR = ενδοτεταρτημοριακό εύρος

² Η θνησιμότητα στη ΜΕΘ περιλαμβάνει θανάτους που συνέβησαν μετά από επανεισαγωγή στη ΜΕΘ.

Πίνακας 8. Σύγκριση κλινικών χαρακτηριστικών και προγνωστικών δεικτών (SAPS II, APACHE II και SOFA) των ασθενών που πήραν εξιτήριο από το νοσοκομείο και εκείνων που απεβίωσαν στο νοσοκομείο

	Επιβίωσαν N=508	Απεβίωσαν N=270	p-value
Φύλο, n(%)			0.22
Θήλυ	175 (62.5)	105 (37.5)	
Άρρεν	333 (66.9)	165 (33.1)	
Ηλικία (έτη) , ΔΜ (IQR)	64.0 (49.0-75.0)	74.0 (64.0-81.0)	<0.001
Ηλικιακή ομάδα ανά τεταρτημόριο (έτη) , n(%)			<0.001
17 - 55	173 (86.5)	27 (13.5)	
55 - 68	131 (63.6)	75 (36.4)	
68 - 77	117 (61.9)	72 (38.1)	
77 - 98	87 (47.5)	96 (52.5)	
Νοσηλείας πριν τη ΜΕΘ (ημέρες) , ΔΜ (IQR)	1.0 (0.0-3.5)	1.0 (0.0-8.0)	0.012
Κυρία διάγνωση εισαγωγής στη ΜΕΘ, n(%)			<0.001
Νόσημα κυκλοφορικού συστήματος	93 (51.4)	88 (48.6)	
Νόσημα αναπνευστικού συστήματος	82 (61.2)	52 (38.8)	
Νεόπλασμα	79 (77.5)	23 (22.5)	
Τραυματισμός ή εξωτερική αιτία	70 (80.5)	17 (19.5)	
Νόσημα πεπτικού συστήματος	45 (66.2)	23 (33.8)	
Λοιμώδες νόσημα	17 (51.5)	16 (48.5)	
Νόσος νευρικού συστήματος	20 (71.4)	8 (28.6)	
Νόσημα ουρογεννητικού συστήματος	13 (86.7)	2 (13.3)	
Άλλη νόσος	89 (68.5)	41 (31.5)	
Ταξινόμηση περιστατικού, n(%)			<0.001
Επείγον χειρουργείο	125 (73.5)	45 (26.5)	
Προγραμματισμένο Χειρουργείο	82 (87.2)	12 (12.8)	
Παθολογικό περιστατικό	226 (53.2)	199 (46.8)	
Τραύμα	75 (84.3)	14 (15.7)	
Πρόελευση περιστατικού, n(%)			0.001
Κλινική/ Άλλη ΜΕΘ	247 (62.1)	151 (37.9)	
ΤΕΠ/Διακομιδή	179 (64.2)	100 (35.8)	
Χειρουργείο	82 (81.2)	19 (18.8)	
SAPS II, MT ± TA	36.6 ± 10.9	48.9 ± 12.8	<0.001
APACHE II, MT ± TA	21.3 ± 6.4	27.9 ± 7.1	<0.001
SOFA, MT ± TA	9.1 ± 3.3	11.5 ± 3.1	<0.001
Συνολική νοσηλεία (ημέρες) , ΔΜ (IQR)	21.0 (12.0-38.0)	22.0 (10.0-38.0)	0.17

Σημειώσεις:

¹ MT = μέση τιμή, TA = τυπική απόκλιση, ΔΜ = διάμεσος, IQR = ενδοτεταρτημοριακό εύρος

4.2. Προγνωστική ακρίβεια της πρόβλεψης θανάτου στη ΜΕΘ

Οι προγνωστικές επιδόσεις των SAPS II, APACHE II και SOFA για την πρόβλεψη της θνησιμότητας στη ΜΕΘ δίνονται στους **Πίνακες 9Α, 9Β, 9Γ**, αντιστοίχως, για ενδεικτικές τιμές των κλιμάκων. Στους πίνακες αυτούς δίνεται και η 'βέλτιστη' τιμή αποκοπής που μεγιστοποιεί το δείκτη Youden και η οποία ήταν 43 για το SAPS II (ευαισθησία Se = 70.6%, ειδικότητα Sp = 65.7%), 26 για το APACHE II (Se = 66.3%, Sp = 71.3%), και 10 για το SOFA (Se = 77.9%, Sp = 50.6%). Κατ' αντιπαράθεση, η ηλικία του ασθενούς έχει βέλτιστη τιμή αποκοπής τα 58 έτη με υψηλή ευαισθησία (Se = 85.3%), αλλά πολύ χαμηλή ειδικότητα (Sp = 31.9%).

Πίνακας 9Α. Προγνωστικές επιδόσεις του SAPS II σε διάφορες ενδεικτικές τιμές και βέλτιστο κατώφλι για την πρόβλεψη θανάτου στη ΜΕΘ

SAPS II	Ευαισθησία (%)	Ειδικότητα (%)	Ακρίβεια (%)	AUC
>= 11	100.0	0.0	21.4	0.50
>= 20	100.0	3.0	23.8	0.52
>= 30	95.1	24.2	39.4	0.60
>= 40	79.1	55.5	60.6	0.67
>= 50	50.3	83.3	76.2	0.67
>= 60	25.2	97.5	82.0	0.61
>= 70	11.0	99.3	80.4	0.55
>= 80	4.9	100.0	79.6	0.52
>= 43 (βέλτιστο*)	70.6	65.7		0.68

*Σημείωση: Το βέλτιστο κατώφλι μεγιστοποιεί το δείκτη Youden (άθροισμα ευαισθησίας και ειδικότητας).

Πίνακας 9B. Προγνωστικές επιδόσεις του APACHE II σε διάφορες ενδεικτικές τιμές και βέλτιστο κατώφλι για την πρόβλεψη θανάτου στη ΜΕΘ

APACHE II	Ευαισθησία (%)	Ειδικότητα (%)	Ακρίβεια (%)	AUC
>= 10	99.4	1.2	22.3	0.50
>= 15	95.7	11.6	29.7	0.54
>= 20	87.1	37.0	47.8	0.62
>= 25	70.6	64.9	66.1	0.68
>= 30	42.3	84.2	75.2	0.63
>= 35	17.8	93.9	77.6	0.56
>= 40	6.7	99.0	79.2	0.53
>= 45	1.8	100.0	78.9	0.51
>= 26 (βέλτιστο*)	66.3	71.3		0.69

*Σημείωση: Το βέλτιστο κατώφλι μεγιστοποιεί το δείκτη Youden (άθροισμα ευαισθησίας και ειδικότητας).

Πίνακας 9Γ. Προγνωστικές επιδόσεις του SOFA σε διάφορες ενδεικτικές τιμές και βέλτιστο κατώφλι για την πρόβλεψη θανάτου στη ΜΕΘ

SOFA	Ευαισθησία (%)	Ειδικότητα (%)	Ακρίβεια (%)	AUC
>= 2	100.0	2.0	23.1	0.51
>= 5	98.8	8.2	27.7	0.54
>= 8	93.9	23.7	38.8	0.59
>= 11	65.0	62.0	62.7	0.64
>= 14	31.9	91.1	78.4	0.61
>= 17	8.0	99.3	79.7	0.54
>= 20	1.2	99.8	78.6	0.51
>= 10 (βέλτιστο*)	77.9	50.6		0.64

*Σημείωση: Το βέλτιστο κατώφλι μεγιστοποιεί το δείκτη Youden (άθροισμα ευαισθησίας και ειδικότητας).

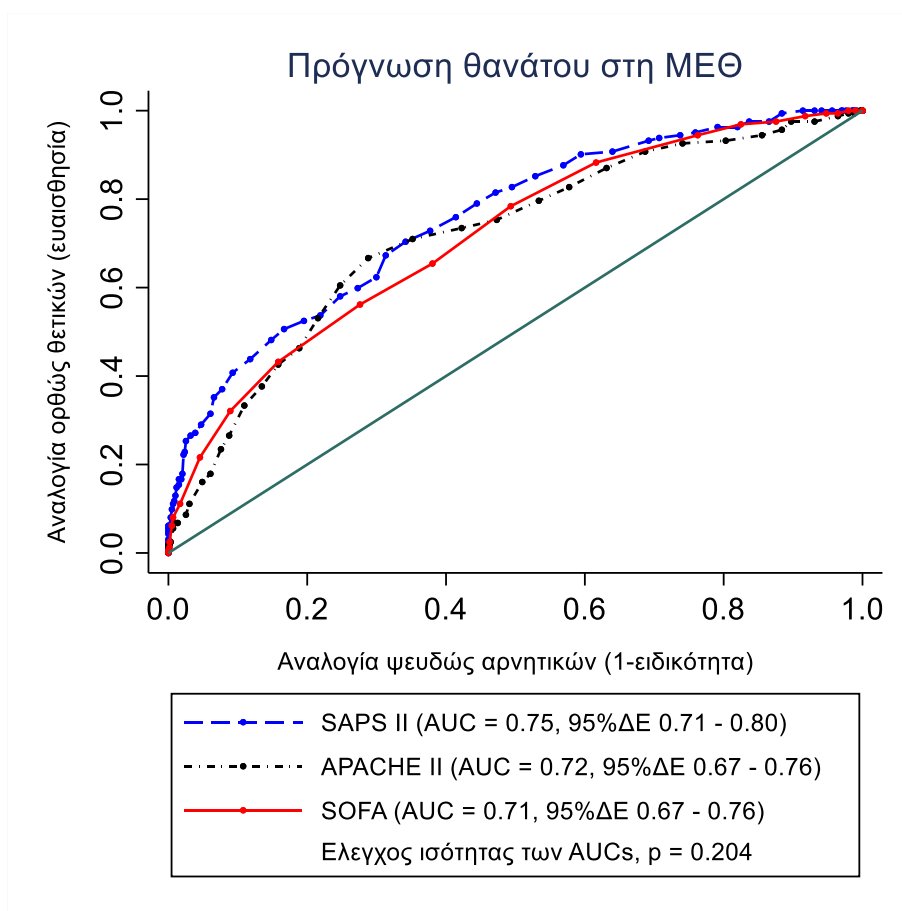
Πίνακας 9Δ. Προγνωστικές επιδόσεις του Ηλικία (έτη) σε διάφορες ενδεικτικές τιμές και βέλτιστο κατώφλι για την πρόβλεψη θανάτου στη ΜΕΘ

Ηλικία (έτη)	Ευαισθησία (%)	Ειδικότητα (%)	Ακρίβεια (%)	AUC
>= 20	100.0	0.2	22.0	0.50
>= 30	98.8	7.6	27.5	0.53
>= 40	94.7	13.7	31.4	0.54
>= 50	91.8	22.0	37.3	0.57
>= 60	80.0	34.5	44.5	0.57
>= 70	55.9	58.7	58.1	0.57
>= 80	27.6	82.1	70.2	0.55
>= 90	0.6	97.5	76.3	0.49
>= 58 (βέλτιστο*)	85.3	31.9		0.59

*Σημείωση: Το βέλτιστο κατώφλι μεγιστοποιεί το δείκτη Youden (άθροισμα ευαισθησίας και ειδικότητας).

Οι καμπύλες ROC, που είναι δείκτες της ικανότητας των προγνωστικών κλιμάκων να διαχωρίζουν μεταξύ ασθενών που επιβιώνουν και εκείνων που πεθαίνουν στη ΜΕΘ, δείχνονται στο **Σχήμα 5**. Τα εμβαδά (AUC) των καμπυλών ROC, ήταν: 0.75 (95%ΔΕ: 0.71 - 0.80) για το SAPS II, 0.72 (95%ΔΕ: 0.67 - 0.76) για το APACHE II, και 0.71 (95%ΔΕ: 0.67) για το SOFA. Παρόλο που η διαχωριστική ικανότητα του SAPS II εμφανίζεται εμφανώς βελτιωμένη σε σύγκριση με τις άλλες δύο κλίμακες, ο στατιστικός έλεγχος ισότητας μεταξύ των εμβαδών ROC [$H_0: AUC(SAPSII) = AUC(APACHEII) = AUC(SOFA)$] δεν έδειξε στατιστικά σημαντική διαφορά ($p = 0.204$).

Σχήμα 5. Καμπύλες ROC της διαχωριστικής ικανότητας των προγνωστικών δεικτών SAPS II, APACHE II και SOFA κατά την πρόβλεψη της θνησιμότητας στη ΜΕΘ

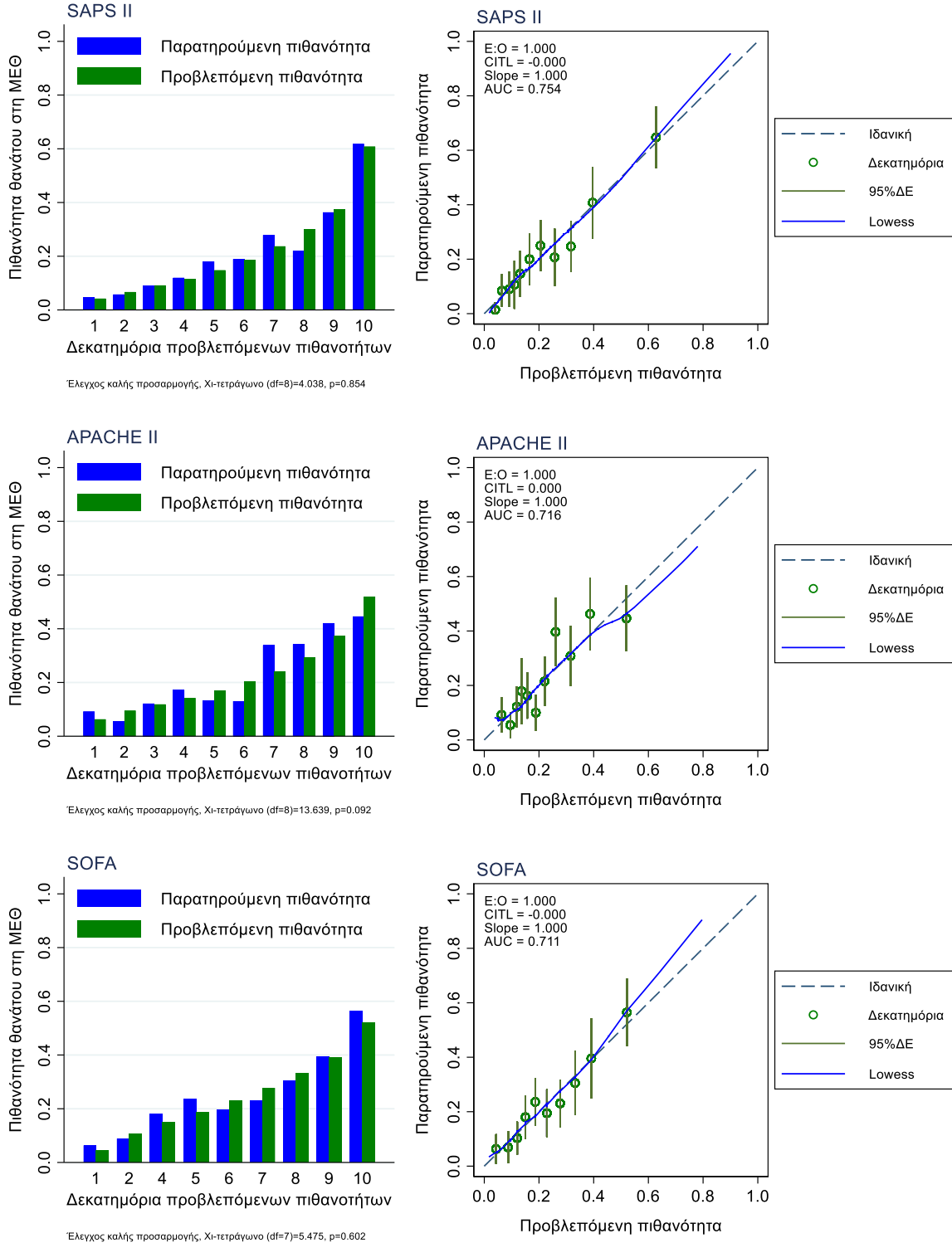


Το **Σχήμα 6** δείχνει τις επιδόσεις βαθμονόμησης των προγνωστικών κλιμάκων SAPS II, APACHE II και SOFA για την πρόβλεψη της θνησιμότητας στη ΜΕΘ.

Στα ραβδογράμματα της αριστερής στήλης, οι ασθενείς χωρίστηκαν σε δέκα παρόμοιες ομάδες κινδύνου που ορίζονται με αύξουσα σειρά της προβλεπόμενης πιθανότητας θανάτου στη ΜΕΘ σύμφωνα με κάθε προγνωστικό μοντέλο και υπολογίστηκαν τα παρατηρούμενα και τα προβλεπόμενα ποσοστά θανάτων σε κάθε ομάδα. Τα ραβδογράμματα βαθμονόμησης κατασκευάστηκαν αντιπαραθέτοντας δίπλα-δίπλα τις προβλεπόμενες και παρατηρούμενες τιμές για κάθε ομάδα και οι μεταξύ τους διαφορές ελέγχθηκαν με τον χι-τετράγωνο έλεγχο καλής προσαρμογής των Hosmer–Lemeshow. Παρατηρούμε ότι είχαμε γενικά πολύ καλή βαθμονόμηση σε όλο το εύρος των πιθανοτήτων θανάτου και ο έλεγχος των Hosmer–Lemeshow δεν εντόπισε στατιστικά σημαντικές διαφορές ανάμεσα στις προβλεπόμενες και τις παρατηρούμενες πιθανότητες ανά δεκατημόριο (αν και για το APACHE II ο έλεγχος ήταν οριακά μη-σημαντικός).

Ωστόσο, τα εξομαλυμένα διαγράμματα βαθμονόμησης (μπλε γραμμή lowess) στη δεξιά στήλη του Σχήματος 6, αποκαλύπτουν ότι το APACHE II και το SOFA έχουν κακή βαθμονόμηση στους ασθενείς υψηλού κινδύνου. Ειδικότερα, το APACHE II φαίνεται να υπερ-εκτιμά τον κίνδυνο θανάτου στη ΜΕΘ όταν η πιθανότητα υπερβαίνει το 40% περίπου, ενώ το SOFA φαίνεται αντιθέτως υπο-εκτιμά τις πιθανότητες θανάτου άνω του 40%. Το SAPS II όμως παρουσίασε εξαιρετικά καλή επίδοση στη βαθμονόμηση.

Σχήμα 6. Διαγράμματα βαθμονόμησης των προγνωστικών δεικτών SAPS II, APACHE II και SOFA για την πρόβλεψη της πιθανότητας θανάτου στη ΜΕΘ



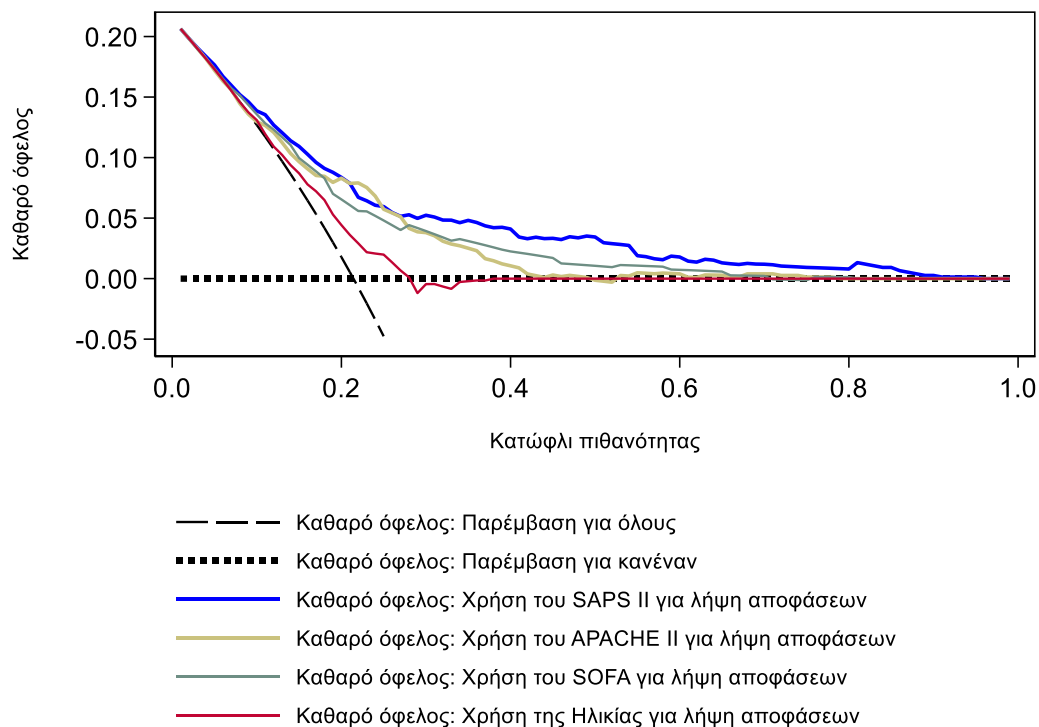
4.3.Καθαρό όφελος και ανάλυση καμπυλών απόφασης

Το **Σχήμα 7** δείχνει τις καμπύλες απόφασης για την κλινική χρησιμότητα και των προγνωστικών κλιμάκων SAPS II, APACHE II και SOFA μετρημένων την πρώτη μέρα εισαγωγής στη ΜΕΘ και της ηλικίας του ασθενούς στην λήψη αποφάσεων με βάση την πρόβλεψη θανάτου στη ΜΕΘ που επιτυγχάνει ο κάθε δείκτης.

Στο γράφημα αυτό, η τετμημένη δείχνει το κατώφλι πιθανότητας, δηλαδή το βαθμό βεβαιότητας για την θνησιμότητα στη ΜΕΘ στον οποίο η απόφαση του ασθενούς ή των γιατρών του θα βασιστεί για τη λήψη παρέμβασης ή αλλαγής στην κλινική διαχείριση (π.χ. εφαρμογή υποστηρικτικής αγωγής ή αποκλιμάκωση της θεραπευτικής αγωγής). Η τεταγμένη του γραφήματος αντιστοιχεί στο καθαρό όφελος από τη χρήση του κάθε προγνωστικού μοντέλου για την πρόβλεψη της πιθανότητας θανάτου στη ΜΕΘ.

Παρατηρούμε ότι οι καμπύλες καθαρού οφέλους των μοντέλων έχουν ελάχιστη επικάλυψη, με το SAPS II να παράγει γενικά μεγαλύτερο καθαρό όφελος συγκριτικά με το APACHE II και το SOFA. Το SAPS II παράγει καθαρό όφελος ανεξάρτητα από το κατώφλι πιθανότητας απόφασης (δηλαδή σχεδόν σε ολόκληρο το εύρος πιθανότητας ,από 0 έως >0.9). Αντιθέτως, η λήψη απόφασης με βάση την ηλικία του ασθενούς δεν έχει όφελος όταν το κατώφλι πιθανότητας είναι πάνω από περίπου 30%.

Σχήμα 7. Καμπύλες απόφασης για την κλινική χρησιμότητα της ηλικίας και των προγνωστικών σκορ SAPS II, APACHE II και SOFA κατά την πρόβλεψη θανάτου στη ΜΕΘ



Οι προβλεπόμενες πιθανότητες θανάτου στη ΜΕΘ σε ενδεικτικές τιμές των κλιμάκων SAPS II, APACHE II, SOFA και της ηλικίας, μαζί με το αντίστοιχο καθαρό όφελος, δίνονται αντιστοίχως στους **Πίνακες 10Α - 10Δ**.

Πίνακας 10Α. Προβλεπόμενη πιθανότητα θανάτου στη ΜΕΘ και καθαρό όφελος του SAPS II

Τιμή SAPS II	Προβλεπόμενη πιθανότητα (%)	95%ΔΕ (%)	Καθαρό όφελος
11	1.8	0.7 - 2.9	0.204
20	3.7	2.0 - 5.4	0.190
30	8.0	5.6 - 10.4	0.157
40	16.4	13.4 - 19.3	0.105
50	30.7	26.5 - 35.0	0.052
60	50.1	42.7 - 57.5	0.035
70	69.4	60.2 - 78.7	0.012
80	83.7	75.7 - 91.8	0.011

Πίνακας 10Β. Προβλεπόμενη πιθανότητα θανάτου στη ΜΕΘ και καθαρό όφελος του APACHE

II

Τιμή APACHE II	Προβλεπόμενη πιθανότητα (%)	95%ΔΕ (%)	Καθαρό όφελος
10	5.1	2.9 - 7.4	0.176
15	8.4	5.8 - 11.1	0.145
20	13.6	10.7 - 16.5	0.113
25	21.3	18.1 - 24.4	0.080
30	31.6	27.2 - 36.0	0.035
35	44.2	37.0 - 51.3	0.001
40	57.5	47.8 - 67.2	0.004
45	69.8	59.0 - 80.7	0.004

Πίνακας 10Γ. Προβλεπόμενη πιθανότητα θανάτου στη ΜΕΘ και καθαρό όφελος του SOFA

Τιμή SOFA	Προβλεπόμενη πιθανότητα (%)	95%ΔΕ (%)	Καθαρό όφελος
2	2.8	1.2 - 4.5	0.196
5	5.9	3.5 - 8.4	0.171
8	12.0	9.1 - 15.0	0.123
11	22.9	19.6 - 26.1	0.054
14	39.1	33.2 - 45.1	0.025
17	58.2	48.5 - 68.0	0.010
20	75.1	64.4 - 85.9	-0.001

Πίνακας 10Δ. Προβλεπόμενη πιθανότητα θανάτου στη ΜΕΘ και καθαρό όφελος της ηλικίας (έτη)

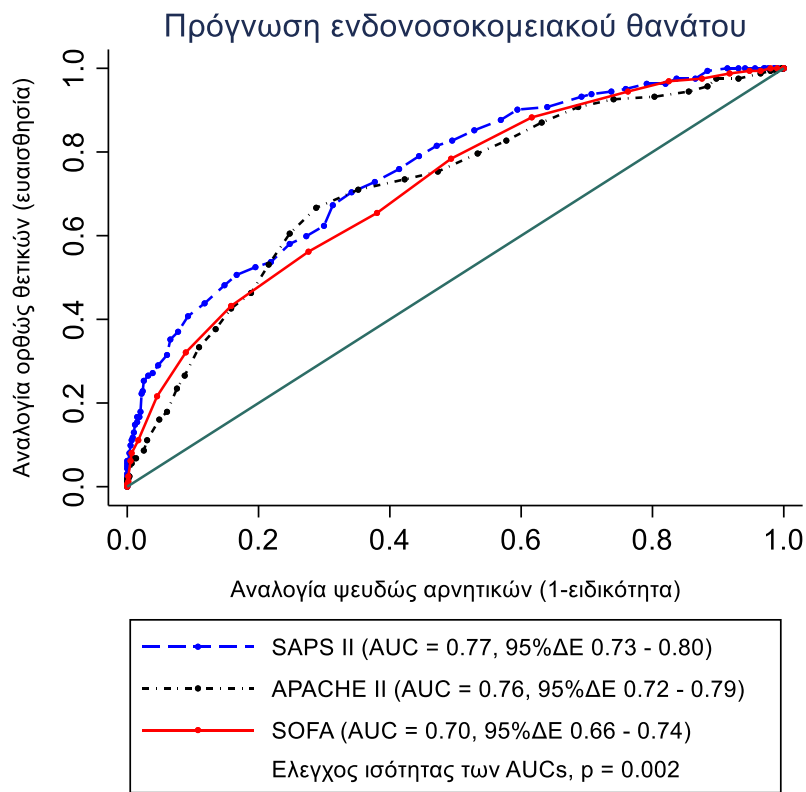
Τιμή Ηλικία (έτη)	Προβλεπόμενη πιθανότητα (%)	95%ΔΕ (%)	Καθαρό όφελος
20	7.6	3.5 - 11.7	0.154
30	9.7	5.6 - 13.8	0.139
40	12.2	8.3 - 16.2	0.113
50	15.3	11.8 - 18.9	0.090
60	19.0	16.0 - 22.1	0.054
70	23.4	20.3 - 26.5	0.023
80	28.4	24.1 - 32.8	0.005
90	34.1	27.4 - 40.8	-0.009

Σημείωση. Η προβλεπόμενη πιθανότητα θανάτου στη ΜΕΘ υπολογίστηκε εισάγωντας κάθε προγνωστικό σκορ (SAPS II, APACHE II, SOFA) ως συνεχή μεταβλητή σε μοντέλο λογιστικής παλινδρόμησης.

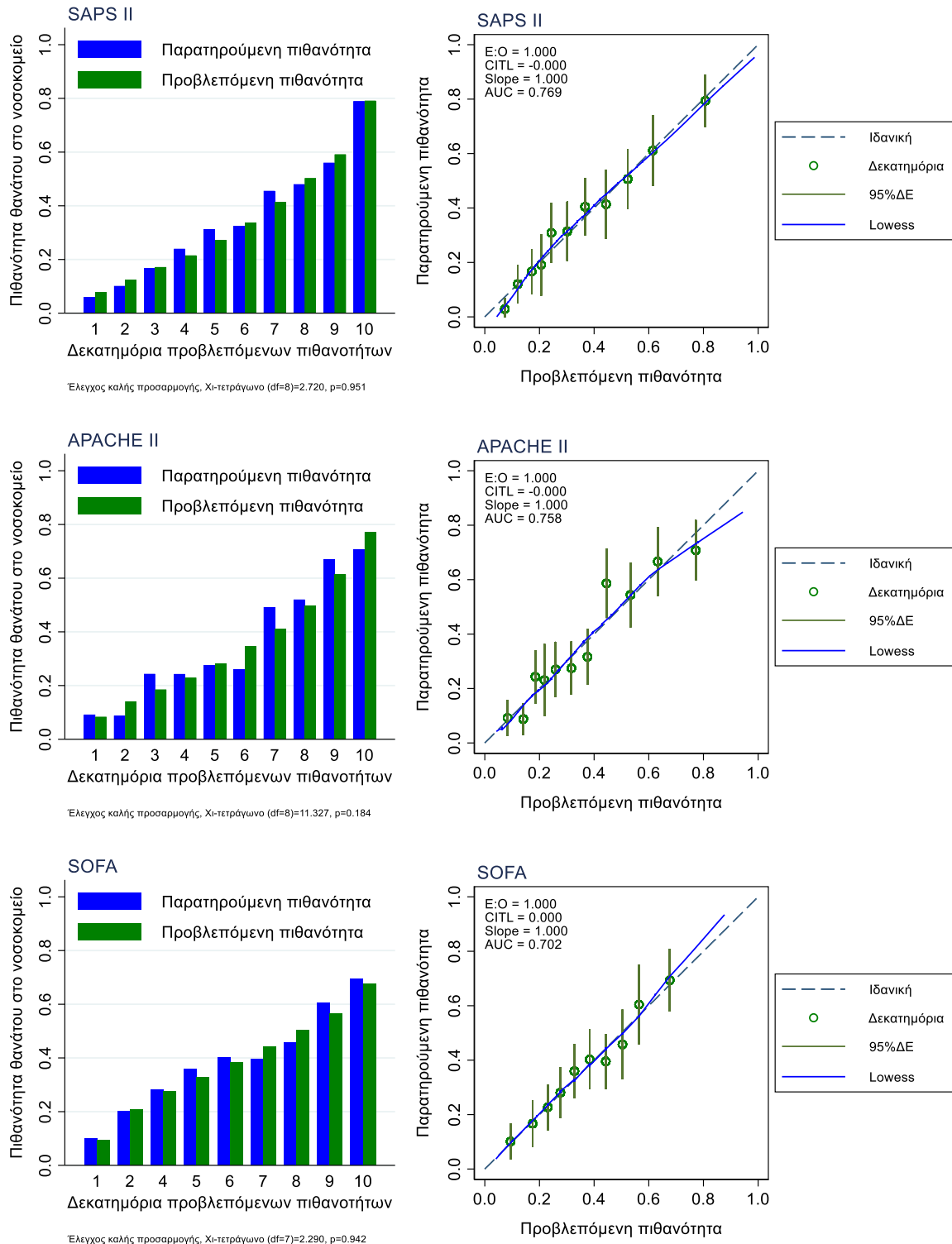
4.4. Πρόβλεψη ενδονοσοκομειακής θνησιμότητας

Επαναλάβαμε τις προηγούμενες αναλύσεις με την έκβαση να είναι ο θάνατος στο νοσοκομείο (εντός της ΜΕΘ ή μετά από την έξοδο από τη ΜΕΘ). Τα αποτελέσματα συνοψίζονται στα **Σχήματα 8,9,10**. Παρατηρούμε ότι η διαχωριστική ικανότητα των SAPS II και APACHE II ήταν ισοδύναμη και στατιστικώς σημαντικά καλύτερη του SOFA για την πρόβλεψη ενδονοσοκομειακού θανάτου σε ασθενείς της ΜΕΘ (Σχήμα 4). Η βαθμονόμηση ήταν καλή για όλα τα προγνωστικά σκορ, και αισθητά βελτιωμένη για τα APACHE II και SOFA για την πρόβλεψη θανάτου στο νοσοκομείο συγκριτικά με την πρόβλεψη θανάτου στη ΜΕΘ (Σχήμα 5). Όσο αφορά το καθαρό όφελος για τη λήψη αποφάσεων, οι επιδόσεις των SAPS II και APACHE II ήταν παρόμοιες και εμφανώς καλύτερες του SOFA. Παρατηρούμε ότι το καθαρό όφελος στο APACHE II αυξήθηκε αισθητά για την πρόγνωση θανάτου στο νοσοκομείο συγκριτικά με την πρόβλεψη θανάτου στη ΜΕΘ.

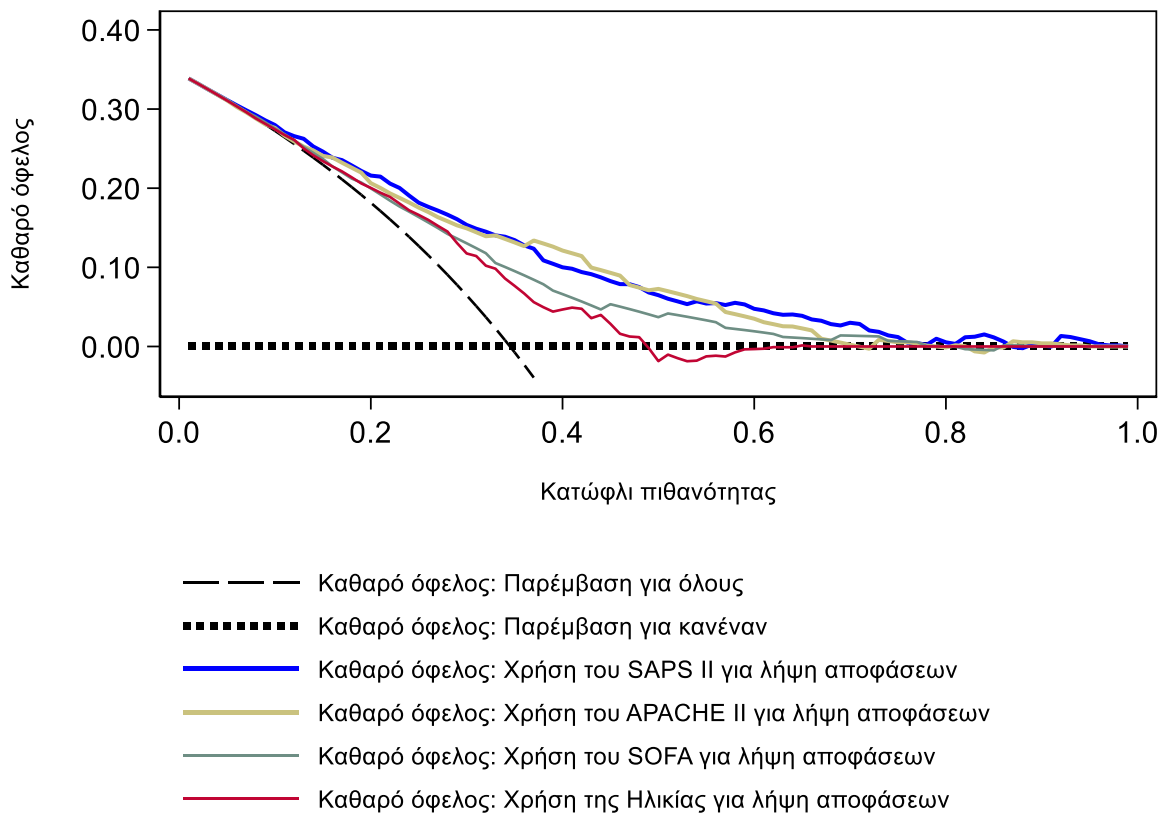
Σχήμα 8. Καμπύλες ROC των προγνωστικών δεικτών SAPS II, APACHE II και SOFA για την πρόβλεψη ενδονοσοκομειακής θνησιμότητας



Σχήμα 9. Διαγράμματα βαθμονόμησης των SAPS II, APACHE II και SOFA για την πρόβλεψη θανάτου στο νοσοκομείο

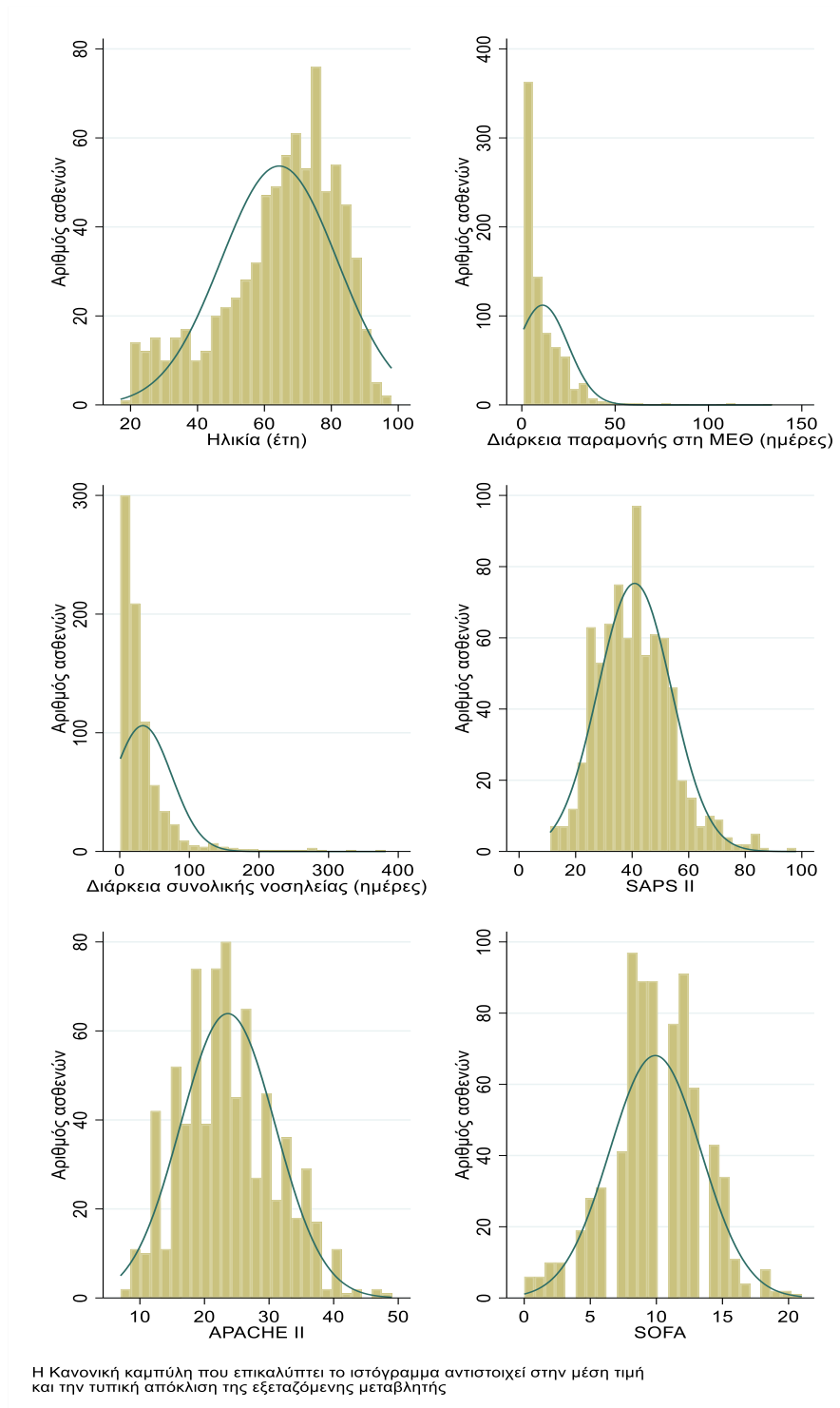


Σχήμα 10. Καμπύλες απόφασης για την κλινική χρησιμότητα της ηλικίας και των προγνωστικών σκορς SAPS II, APACHE II και SOFA κατά την πρόβλεψη ενδονοσοκομειακού θανάτου σε ασθενείς της ΜΕΘ

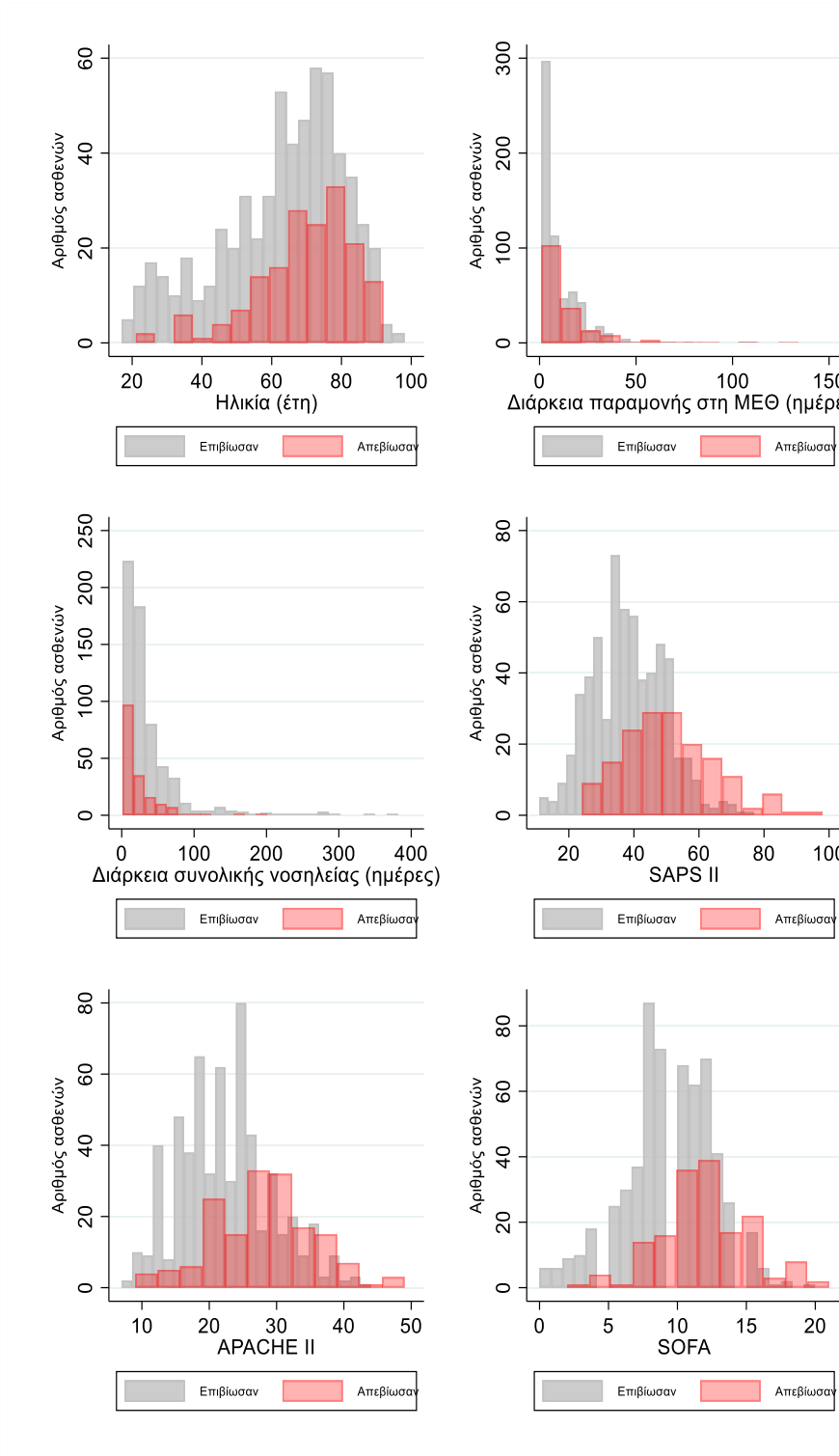


4.5. Παράρτημα συμπληρωματικών διαγραμμάτων

Συμπληρωματικό σχήμα S1. Κατανομές των συνεχών μεταβλητών της μελέτης.



Συμπληρωματικό σχήμα S2. Σύγκριση των κατανομών συνεχών μεταβλητών της μελέτης ανάμεσα σε ασθενείς που επιβίωσαν και ασθενείς που απεβίωσαν στη ΜΕΘ



5. Συζήτηση

Στην παρούσα μελέτη, έγινε μια προσπάθεια εξωτερικής επικύρωσης και αξιολόγησης της απόδοσης τριών ευρέως χρησιμοποιούμενων προγνωστικών μοντέλων (SAPS II, APACHE II και SOFA), σχετικά με τη δυνατότητα πρόβλεψης της θνησιμότητας εντός της Μονάδας Εντατικής Θεραπείας του ΠαΓΝΗ και του νοσοκομείου γενικότερα. Τα δεδομένα προήλθαν από βαρέως πάσχοντες ασθενείς με διάφορες αιτίες εισαγωγής (χειρουργικά, παθολογικά, τραυματίες) που εισήχθησαν για εντατική παρακολούθηση και θεραπεία. Τα μοντέλα δοκιμάστηκαν στο αν θα μπορούσαν να βοηθήσουν στη λήψη κλινικών αποφάσεων σύμφωνα με την ικανότητα πρόβλεψης της έκβασης των περιστατικών και του κινδύνου θανάτου. Από τους ηλεκτρονικούς ιατρικούς φακέλους των ασθενών (βάσεις CRITIS και AMED), συλλέχθηκαν μεταβλητές (κλινικές, εργαστηριακές, δημογραφικές) οι οποίες συνδυαστικά κατέληξαν στο τελικό αθροιστικό σκορ για κάθε ασθενή και για κάθε προγνωστικό, σύμφωνα με τις πρωτότυπες περιγραφές των μοντέλων και χωρίς να γίνουν προσαρμογές στις επιμέρους μεταβλητές ή τη στάθμιση τους.

Όπως αναφέρεται και σε παρόμοιες μελέτες, το SAPS II και το APACHE II αναπτύχθηκαν αρχικά με σκοπό να προβλέψουν την ενδονοσοκομειακή θνησιμότητα, ενώ το SOFA για μέτρηση της οργανικής δυσλειτουργίας και κατόπιν επικυρώθηκε για την συμβολή του στην πρόβλεψη θνησιμότητας σε διαφορετικές ομάδες ασθενών^{28,29,8,11}. Και σε αντίστοιχες προσπάθειες ανάλυσης καμπυλών ROC σε άλλες χώρες, έχουν προκύψει αποτελέσματα από καμπύλες ROC (AUC) που κυμαίνονται μεταξύ 0,7 και 0,9, για τα SAPS II και APACHE II, σχετικά με την ενδονοσοκομειακή θνησιμότητα¹⁵.

Αξίζει να αναφερθεί ότι το δείγμα της παρούσας μελέτης ήταν ασθενείς μικτής Μονάδας Εντατικής Θεραπείας ενηλίκων, ενός τριτοβάθμιου πανεπιστημιακού νοσοκομείου αστικής περιοχής. Στα μοντέλα που αναλύθηκαν για το δείγμα αυτό δεν έχει υπάρξει προηγούμενη εξωτερική επικύρωση της προβλεπτικής ακρίβειας. Σύμφωνα με τα αποτελέσματα που της παρούσας μελέτης, τα μοντέλα SAPS II, APACHE II και SOFA έχουν ικανοποιητικά επίπεδα διαχωριστικής ικανότητας και βαθμονόμησης στην πρόγνωση της ενδονοσοκομειακής θνησιμότητας. Τα επίπεδα αυτά επιτυγχάνονται κατά ένα βαθμό παραπάνω από το SAPS II, τόσο στην παρούσα μελέτη όσο και σε αντίστοιχη από τη διεθνή βιβλιογραφία. Για την πρόγνωση της ενδονοσοκομειακής θνησιμότητας, φάνηκε ότι και το APACHE II σημείωσε καλύτερη απόδοση από το SOFA, το οποίο είχε την τάση να υποτιμά τους ασθενείς χαμηλού κινδύνου και να υπερεκτιμά τους ασθενείς υψηλού, σχετικά με την πιθανότητα θανάτου¹¹.

Σε άλλη μελέτη, το SAPS II έδειξε να έχει κατά ένα βαθμό λιγότερη ακρίβεια στην πρόβλεψη πιθανότητας θανάτου εντός της ΜΕΘ. Το αποτέλεσμα αυτό έρχεται σε διαφωνία με την παρούσα μελέτη, όπου αναδείχθηκε ότι η εν λόγω κλίμακα εμφανίζει την καλύτερη βαθμονόμηση σε ολόκληρο το φάσμα πιθανοτήτων θανάτου, τόσο εντός της ΜΕΘ όσο και εντός του νοσοκομείου γενικότερα. Το SAPS II, όπως και το APACHE II ενδέχεται γενικά να έχουν καλύτερη προβλεπτική επίδοση για την πρόβλεψη θανάτου εντός ΜΕΘ, λόγω του ότι, για να υπολογιστούν, χρησιμοποιούν χαρακτηριστικά που αντικατοπτρίζουν χρόνιες καταστάσεις, όπως η ηλικία του ασθενούς ή η αιτία εισαγωγής του, εν αντιθέσει με το SOFA, το οποίο λαμβάνει υπόψιν έξι μεταβλητές που αντανακλούν ως επί το πλείστον οξείες μεταβολές. Παρόλο που η βαθμονόμηση μεταξύ παρατηρούμενων και προβλεπόμενων πιθανοτήτων ήταν πολύ καλή και για τα τρία προγνωστικά μοντέλα για την πρόβλεψη ενδονοσοκομειακού θανάτου (εντός ή εκτός ΜΕΘ), παρατηρήσαμε κακή βαθμονόμηση των APACHE II και SOFA για την πρόβλεψη θανάτου εντός της ΜΕΘ για τους ασθενείς υψηλού κινδύνου. Ειδικότερα, το APACHE II φάνηκε να υπερ-εκτιμά τον κίνδυνο θανάτου στη ΜΕΘ όταν η πιθανότητα υπερβαίνει το 40% περίπου, ενώ το SOFA φάνηκε αντιθέτως υπο-εκτιμά τις πιθανότητες θανάτου άνω του 40%. Το SAPS II όμως παρουσίασε εξαιρετικά καλή επίδοση στη βαθμονόμηση. Η επαναβαθμονόμηση (recalibration) ενός προγνωστικού μοντέλου μπορεί να επιφέρει βελτίωση όταν ο πληθυσμός στον οποίο αξιολογείται το μοντέλο διαφέρει πολύ από τον πληθυσμό στον οποίο το μοντέλο αναπτύχθηκε (π.χ. διαφορετική θνησιμότητα). Μελλοντική ανάλυση των δεδομένων της μελέτης θα πρέπει επομένως να συγκρίνει τα χαρακτηριστικά των ασθενών μελέτης με εκείνα του πληθυσμού στον οποίο αναπτύχθηκαν τα σκορς APACHE II και SOFA και να διερευνήσει αν μια επαναβαθμονόμηση των μοντέλων αυτών είναι απαραίτητη για την ακριβή πρόβλεψη θανάτου στη ΜΕΘ στους ασθενείς του δικού μας νοσοκομείου¹¹.

Η παραδοσιακή στατιστική αξιολόγηση των προγνωστικών μοντέλων, χρησιμοποιώντας μέτρα διαχωριστικής ικανότητας και βαθμονόμησης δεν μπορεί να εξασφαλίσει με σιγουριά εάν ένα συγκεκριμένο μοντέλο αξίζει να εφαρμοστεί ως βοηθητικό εργαλείο στη λήψη κλινικών αποφάσεων ή ποιο από αυτά είναι το καλύτερο στην πράξη^{30,31}. Οπότε, εδώ είναι που κρίνεται ωφέλιμη η πραγματοποίηση της ανάλυσης καθαρού οφέλους^{32,33}.

Στην παρούσα μελέτη, εξετάσαμε στα δοκιμασμένα αυτά μοντέλα το καθαρό όφελος, σε πολλά διαφορετικά «κατώφλια» κινδύνου. Αδιαμφισβήτητα, ο θάνατος είναι ένα σοβαρό τελικό γεγονός, το οποίο προσπαθεί να αποτραπεί από τη στιγμή που ένας ασθενής εισάγεται στη ΜΕΘ. Τα ψευδή

αποτελέσματα που μπορεί να προέλθουν από τον υπολογισμό βαθμολογιών που μπορεί να μην είναι τόσο ακριβής, τελικά περιορίζουν την σωστή εφαρμογή τους και έτσι δημιουργούνται ασάφειες στην πρόγνωση. Αυτό εν συνεχεία οδηγεί σε αμφιβολία σχετικά με το όφελος πλήρους θεραπευτικής επένδυσης σε ασθενείς με αυξημένο κίνδυνο θνησιμότητας και εν τέλει επιφέρει πρόσθετη ταλαιπωρία και περιττή χρησιμοποίηση των πόρων^{30,34}.

Τα δεδομένα που προέκυψαν, όπως και σε αντίστοιχες έρευνες, υποδηλώνουν ότι τα SAPS II, APACHE II και SOFA που υπολογίζονται κατά την εισαγωγή του ασθενούς μπορεί να προσθέσουν πρόσθετες πληροφορίες που θα συμβάλουν στη λήψη αποφάσεων σχετικά με τη θεραπευτική διαχείριση και την παρηγορητική φροντίδα, όταν ο προβλεπόμενος κίνδυνος μπορεί να είναι εντός συγκεκριμένων ορίων. Αυτό είναι πιο εύκολο να εφαρμοστεί σε ασθενείς με υψηλό κίνδυνο θανάτου, που έχει και περισσότερη σημασία. Σε ασθενείς με επίπεδα σοβαρότητας χαμηλά ή ενδιάμεσα είναι πιο δύσκολο να προβλεφθεί ο κίνδυνος θανάτου. Η μέτρηση των βαθμολογιών καθ' όλη τη διάρκεια της νοσηλείας, σε συνδυασμό και με όλα τα υπόλοιπα στοιχεία που αφορούν τον ασθενή μπορεί να συμβάλει στο σχεδιασμό ενός καλύτερου πλάνου φροντίδας και να βελτιώσει τις αποφάσεις σχετικά με τη θεραπευτική διαχείριση^{34,30,35}.

Για τη μελέτη αυτή δεν υπήρξαν κριτήρια αποκλεισμού. Εξαιρέθηκε από τις αναλύσεις ένας μικρός αριθμός ασθενών (20 άτομα, 2.6%) για τους οποίους ορισμένες μεταβλητές δεν είχαν καταγραφεί, οπότε δεν μπορούσε να υπολογιστεί η τελική βαθμολογία. Το δείγμα προήλθε από ένα ενιαίο σύνολο, χωρίς να υπάρχει κάποιο συγκεκριμένο κριτήριο επιλογής και ταυτόχρονα έδωσε μια σημαντική εικόνα σχετικά με τους λόγους εισαγωγής των ασθενών στη ΜΕΘ, το διάστημα παραμονής τους εκεί και στο νοσοκομείο και τις διαφοροποιήσεις ανάλογα με το φύλο και την ηλικία. Η θνησιμότητα εντός της ΜΕΘ σημειώθηκε κοντά στο 22% και εντός του νοσοκομείου κοντά στο 35%. Σε άλλη μελέτη, όπου οι ασθενείς ανήκαν σε μια συγκεκριμένη κατηγορία (χειρουργικοί ασθενείς), τα ανάλογα ποσοστά ήταν κάτω του 10%, πιθανόν λόγω της διαφοράς αυτής¹¹.

Αξίζει να σημειωθεί ότι, η χρονική στιγμή που πραγματοποιούνται οι υπολογισμοί της κάθε κλίμακας/σκορ και η χρονική διαφορά μεταξύ τους δημιουργεί μια ποικιλία χαρακτηριστικών στις θεραπευτικές επιλογές, καθώς και στο αντίκτυπο που θα έχουν στη λήψη αποφάσεων. Η βάση δεδομένων του SAPS χτίστηκε σύμφωνα με τα δεδομένα ασθενών που εισήχθησαν σε Μονάδες Εντατικής Θεραπείας έως και το 2002, ενώ αντίστοιχα το APACHE II συμπεριέλαβε ασθενείς από το 1979 έως το 1982, σε πολλαπλές ΜΕΘ των ΗΠΑ^{29,28}. Τα δεδομένα της παρούσας μελέτης εδώ

προέρχονταν από το χρονικό διάστημα 2018-2020. Οι διαφορές στη συχνότητα των νοσημάτων και των εκβάσεων τους από τότε έως τώρα αποτελούν ένα σημαντικό χαρακτηριστικό που μπορεί να επηρεάσει τα αποτελέσματα που προκύπτουν σε μελέτες εξωτερικής επικύρωσης¹¹.

Η αντιμετώπιση και οι αποφάσεις θεραπείας των ασθενών σε Μονάδες Εντατικής Θεραπείας προκαλεί όπως φάνηκε προβληματισμό, σχετικά με την κατεύθυνση που πρέπει να ακολουθηθεί. Υπάρχει πλέον η δυνατότητα σύνθετων επιλογών που θα έχουν στόχο το καλύτερο δυνατό αποτέλεσμα, σύμφωνα με την πρόγνωση του περιστατικού^{36,37}. Σε πρόσφατη μελέτη αναφέρθηκε ότι, στο 14% των ασθενών που θα εισαχθούν σε μια ΜΕΘ, μετά από κάποιο χρονικό διάστημα θα αποφασιστεί η απόσυρση ή η διακοπή της θεραπείας³⁸. Εάν οι συνθήκες ήταν ιδανικές, ο ασθενής θα ήταν αυτός που θα καθόριζε την εξέλιξη, επομένως οι επαγγελματίες υγείας κρίνονται τελικά οι αρμόδιοι.

Οι αποφάσεις προς το τέλος του κύκλου ζωής έχει φανεί ότι μεταβάλλονται εύκολα, σύμφωνα με αρκετές μελέτες σε διάφορες χώρες^{39,40,41,42}. Γίνεται λοιπόν σαφές ότι, η ύπαρξη μιας αξιόπιστης προγνωστικής κλίμακας, με όσο το δυνατόν υψηλή ευαισθησία και ειδικότητα, μπορεί να παίξει σημαντικό ρόλο σε αυτό τον τομέα. Η ακριβής αξιολόγηση της αρχικής κλινικής κατάστασης, των διαδικασιών που θα ακολουθήσουν και του τελικού αποτελέσματος είναι ζωτικής σημασίας, εκτιμώντας έτσι των κίνδυνο επιπλοκών αλλά και τη βέλτιστη απόδοση των θεραπευτικών προσπαθειών¹¹.

Δεν πρέπει να παραλείπεται το γεγονός ότι, ένα μεγάλο δείγμα ασθενών, προερχόμενο από διάφορες κατηγορίες για εξακρίβωση της απόδοσης των προγνωστικών αυτών μοντέλων συμβάλουν σημαντικά στο τελικό αποτέλεσμα. Η εξωτερική επικύρωση δεν πρέπει να παραλείπεται με σκοπό την αξιολόγηση της ακρίβειας σε νέες υποομάδες ασθενών και σε διαφορετικές συνθήκες ανάλογα με το χρόνο. Είναι προφανές όμως, ότι οι κλίμακες βαρύτητας από μόνες τους δεν μπορούν ποτέ να αντικαταστήσουν πλήρως την ανθρώπινη λήψη αποφάσεων. Για να συνδυαστούν σωστά, χρειάζονται αξιόπιστη και συνεχής συλλογή πληροφοριών.

6. Συμπέρασμα

Σε αυτή τη μελέτη παρατηρήθηκε ότι, οι προγνωστικές βαθμολογίες SAPS II, Apache II και SOFA είναι επαρκή εργαλεία για την πρόβλεψη θνησιμότητας εντός της ΜΕΘ και εντός του νοσοκομείου σε μια ομάδα ασθενών από μια μικτή ΜΕΘ στην Ελλάδα. Το SAPS II ήταν το μοντέλο καλύτερης απόδοσης, τόσο για τη θνησιμότητα εντός της ΜΕΘ όσο και για τη νοσοκομειακή θνησιμότητα, αλλά το APACHE II θα μπορούσε επίσης να είναι χρήσιμο για την αναγνώριση των ασθενών που κινδυνεύουν να πεθάνουν στο νοσοκομείο. Όλες οι βαθμολογίες, ειδικά το SAPS II, έδειξαν δυνατότητες να χρησιμοποιηθούν ως βοηθητικά εργαλεία για τη στήριξη της λήψης αποφάσεων από τους ιατρούς και τις οικογένειες σχετικά με το επίπεδο της θεραπευτικής διαχείρισης και της παρηγορητικής φροντίδας.

Όπως έχει προταθεί και από παρόμοιες μελέτες, τα δεδομένα εδώ υποστηρίζουν την υιοθέτηση και εφαρμογή των προγνωστικών αυτών βαθμολογιών στην καθημερινή κλινική πράξη, πραγματοποιώντας όταν χρειάζεται εξωτερική επικύρωση. Αξίζει να σημειωθεί ότι κανένα προγνωστικό μοντέλο δεν πρέπει να χρησιμοποιείται μεμονωμένα για να καθοδηγήσει τη λήψη αποφάσεων ή για να αντικαταστήσει την κλινική κρίση. Περαιτέρω μελέτες θα βοηθούσαν στον ακριβή καθορισμό του ρόλου που μπορεί να έχουν τα δοκιμασμένα μοντέλα ως μέρος της διαδικασίας λήψης αποφάσεων στη ΜΕΘ.

7. Βιβλιογραφία

1. Νάκος, Γ., Καραχάλιου, Α., Κουλούρας, Β. & Κωνσταντή, Ε. *Εντατική θεραπεία - Αρχές και Εξελίξεις*. (Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, 2015).
2. Κουτσούκου, Α. *Βασικές Αρχές Εντατικής Θεραπείας*. (Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα, 2015).
3. Φίκα, Σ., Νανάς, Σ., Μπαλτόπουλος, Γ. & Μυριανθεύς, Π. *Προγνωστικά συστήματα και προγνωστικοί δείκτες έκβασης ασθενών της μονάδας εντατικής θεραπείας*. vol. 5 (Αρχεία Ελληνικής Ιατρικής, 2014).
4. Zimmerman, J. E., Kramer, A. A., McNair, D. S., Malila, F. M. & Shaffer, V. L. Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV*. *Crit. Care Med.* **34**, 2517–2529 (2006).
5. Gall, J.-R. L. *et al.* A simplified acute physiology score for ICU patients: *Crit. Care Med.* **12**, 975–977 (1984).
6. Le Gall, J. R. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA J. Am. Med. Assoc.* **270**, 2957–2963 (1993).
7. Moreno, R. P. *et al.* SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med.* **31**, 1345–1355 (2005).
8. Vincent, J.-L. *et al.* The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure: On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine (see contributors to the project in the appendix). *Intensive Care Med.* **22**, 707–710 (1996).
9. Ferreira, F. L. Serial Evaluation of the SOFA Score to Predict Outcome in Critically Ill Patients. *JAMA* **286**, 1754 (2001).
10. Keegan, M. T. & Soares, M. What every intensivist should know about prognostic scoring systems and risk-adjusted mortality. *Rev. Bras. Ter. Intensiva* **28**, 264–269 (2016).
11. Falcão, A. L. E. *et al.* The prognostic accuracy evaluation of SAPS 3, SOFA and APACHE II scores for mortality prediction in the surgical ICU: an external validation study and decision-making analysis. *Ann. Intensive Care* **9**, 18 (2019).
12. Fuchs, P. A., Czech, I. J. & Krzych, Ł. J. The Pros and Cons of the Prediction Game: The Never-ending Debate of Mortality in the Intensive Care Unit. *Int. J. Environ. Res. Public Health* **16**, E3394 (2019).
13. Schoe, A., Bakhshi-Raiez, F., de Keizer, N., van Dissel, J. T. & de Jonge, E. Mortality prediction by SOFA score in ICU-patients after cardiac surgery; comparison with traditional prognostic-models. *BMC Anesthesiol.* **20**, 65 (2020).

14. Koch, C. *et al.* Comparison of qSOFA score, SOFA score, and SIRS criteria for the prediction of infection and mortality among surgical intermediate and intensive care patients. *World J. Emerg. Surg. WJES* **15**, 63 (2020).
15. Sarkar, R. *et al.* Performance of intensive care unit severity scoring systems across different ethnicities in the USA: a retrospective observational study. *Lancet Digit. Health* **3**, e241–e249 (2021).
16. Allyn, J. *et al.* Simplified Acute Physiology Score II as Predictor of Mortality in Intensive Care Units: A Decision Curve Analysis. *PloS One* **11**, e0164828 (2016).
17. Kranidiotis, G. *et al.* End-of-life decisions in Greek intensive care units: a multicenter cohort study. *Crit. Care Lond. Engl.* **14**, R228 (2010).
18. Vickers, A. J. & Elkin, E. B. Decision curve analysis: a novel method for evaluating prediction models. *Med. Decis. Mak. Int. J. Soc. Med. Decis. Mak.* **26**, 565–574 (2006).
19. Zhang, Z. *et al.* Decision curve analysis: a technical note. *Ann. Transl. Med.* **6**, 308 (2018).
20. Debray, T. P. A. *et al.* A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J. Clin. Epidemiol.* **68**, 279–289 (2015).
21. Habibzadeh, F., Habibzadeh, P. & Yadollahie, M. On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem. Medica* **26**, 297–307 (2016).
22. Van Calster, B. *et al.* Calibration: the Achilles heel of predictive analytics. *BMC Med.* **17**, 230 (2019).
23. Van Calster, B. *et al.* Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur. Urol.* **74**, 796–804 (2018).
24. Detering, K. M., Hancock, A. D., Reade, M. C. & Silvester, W. The impact of advance care planning on end of life care in elderly patients: randomised controlled trial. *BMJ* **340**, c1345 (2010).
25. Scheunemann, L. P., McDevitt, M., Carson, S. S. & Hanson, L. C. Randomized, controlled trials of interventions to improve communication in intensive care: a systematic review. *Chest* **139**, 543–554 (2011).
26. Kon, A. A. *et al.* Shared Decision Making in ICUs: An American College of Critical Care Medicine and American Thoracic Society Policy Statement. *Crit. Care Med.* **44**, 188–201 (2016).
27. Pauker, S. G. & Kassirer, J. P. Therapeutic decision making: a cost-benefit analysis. *N. Engl. J. Med.* **293**, 229–234 (1975).
28. Metnitz, P. G. H. *et al.* SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. *Intensive Care Med.* **31**, 1336–1344 (2005).
29. Knaus, W. A., Draper, E. A., Wagner, D. P. & Zimmerman, J. E. APACHE II: a severity of disease classification system. *Crit. Care Med.* **13**, 818–829 (1985).

30. Vickers, A. J. & Elkin, E. B. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med. Decis. Making* **26**, 565–574 (2006).
31. Alba, A. C. *et al.* Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. *JAMA* **318**, 1377 (2017).
32. Salluh, J. I. F. & Soares, M. ICU severity of illness scores: APACHE, SAPS and MPM. *Curr. Opin. Crit. Care* **20**, 557–565 (2014).
33. Vincent, J.-L. & Moreno, R. Clinical review: Scoring systems in the critically ill. *Crit. Care* **14**, 207 (2010).
34. Vickers, A. J. Decision analysis for the evaluation of diagnostic tests, prediction models and molecular markers. *Am. Stat.* **62**, 314–320 (2008).
35. Schenker, Y. *et al.* "It Hurts To Know...And It Helps": Exploring How Surrogates in the ICU Cope with Prognostic Information. *J. Palliat. Med.* **16**, 243–249 (2013).
36. Truog, R. D. *et al.* Recommendations for end-of-life care in the intensive care unit: a consensus statement by the American College [corrected] of Critical Care Medicine. *Crit. Care Med.* **36**, 953–963 (2008).
37. Cook, D. & Rocker, G. Dying with dignity in the intensive care unit. *N. Engl. J. Med.* **370**, 2506–2514 (2014).
38. Lesieur, O., Leloup, M., Gonzalez, F., Mamzer, M.-F., & EPILAT study group. Withholding or withdrawal of treatment under French rules: a study performed in 43 intensive care units. *Ann. Intensive Care* **5**, 56 (2015).
39. Wilkinson, D. J. C. & Truog, R. D. The luck of the draw: physician-related variability in end-of-life decision-making in intensive care. *Intensive Care Med.* **39**, 1128–1132 (2013).
40. Luce, J. M. End-of-life decision making in the intensive care unit. *Am. J. Respir. Crit. Care Med.* **182**, 6–11 (2010).
41. Phua, J. *et al.* Withholding and withdrawal of life-sustaining treatments in intensive care units in Asia. *JAMA Intern. Med.* **175**, 363–371 (2015).
42. Mark, N. M., Rayner, S. G., Lee, N. J. & Curtis, J. R. Global variability in withholding and withdrawal of life-sustaining treatment in the intensive care unit: a systematic review. *Intensive Care Med.* **41**, 1572–1585 (2015).