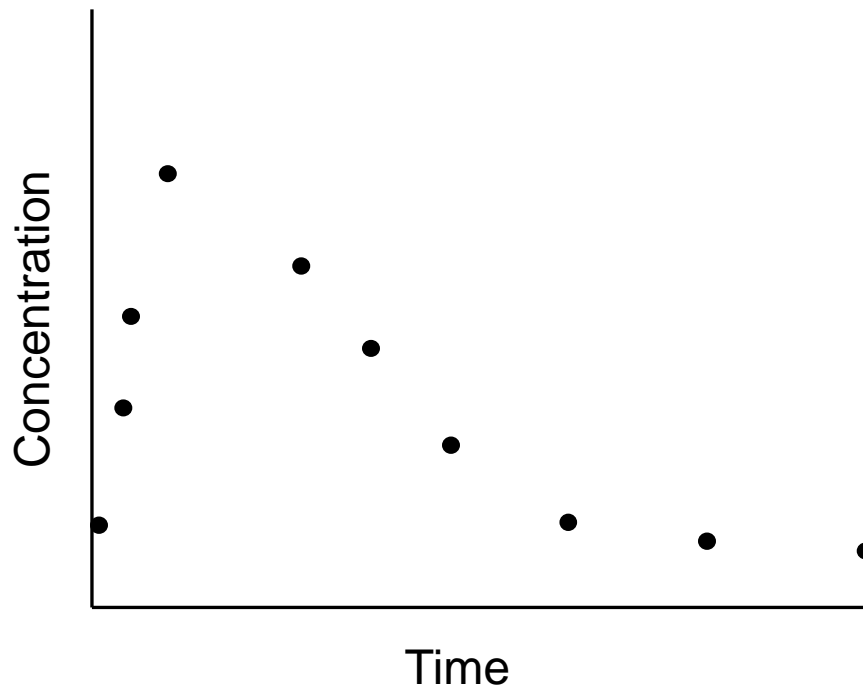
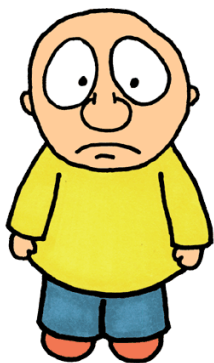


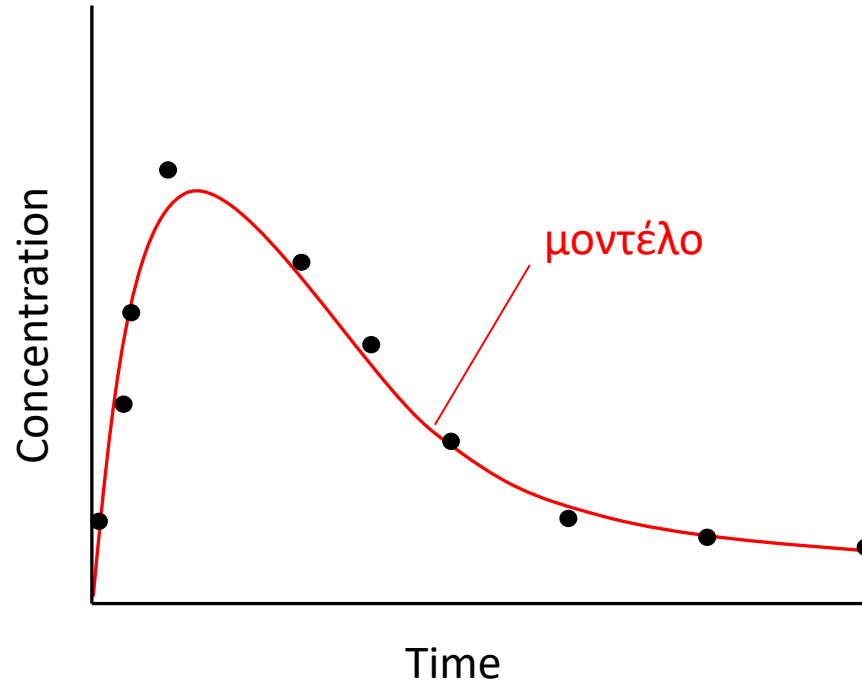
Πληθυσμιακή ΦΚ – Nonlinear mixed effects models

Πληθυσμιακή ΦΚ – Nonlinear mixed effects models

Φαρμακοκινητικό προφίλ ενός ασθενούς



Χρησιμοποιούνται παραμετρικά μαθηματικά μοντέλα



παράμετροι CL , V , k_a

$$C = \frac{D \cdot k_a}{k_a \cdot V - CL} \left(e^{-\frac{CL}{V}t} + e^{-k_a t} \right)$$

Εκτίμηση παραμέτρων για έναν ασθενή με μη γραμμική παλινδρόμηση

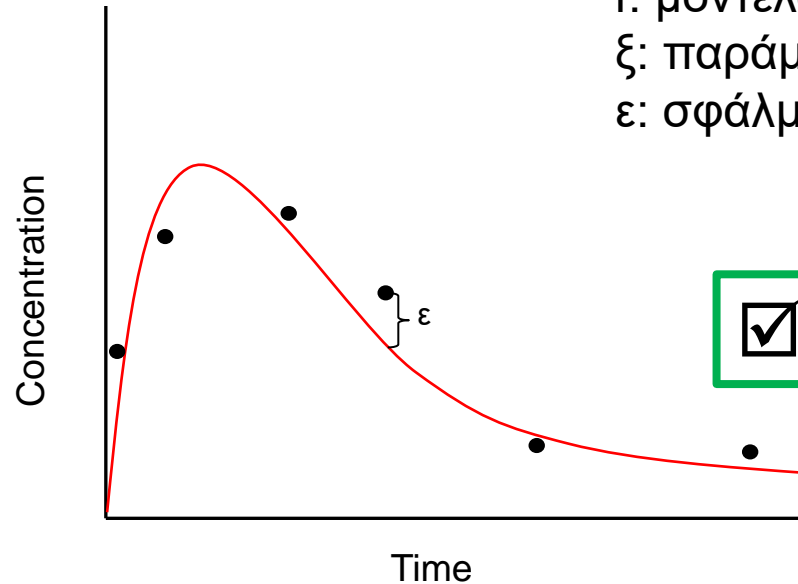
$$C_j = f(\xi, t_j) + \varepsilon_j$$

C: μετρήσεις

f: μοντέλο

ξ : παράμετροι (CL, V, k_a)

ε : σφάλματα



Εκτίμηση παραμέτρων

Στόχος : να εκτιμηθούν οι παράμετροι ξ έτσι ώστε να ελαχιστοποιούνται Οι αποστάσεις (σφάλματα) ε_j του μοντέλου **f** από τις μετρήσεις **C**.

Τα σφάλματα ε_j οφείλονται σε:

σφάλματα στις μετρήσεις
σφάλματα στους χρόνους δειγματοληψίας
ατελή συμφωνία του μοντέλου
(misspecification)

Μέθοδος ελαχίστων τετραγώνων

Θέλουμε να ελαχιστοποιήσουμε τα σφάλματα ε_j

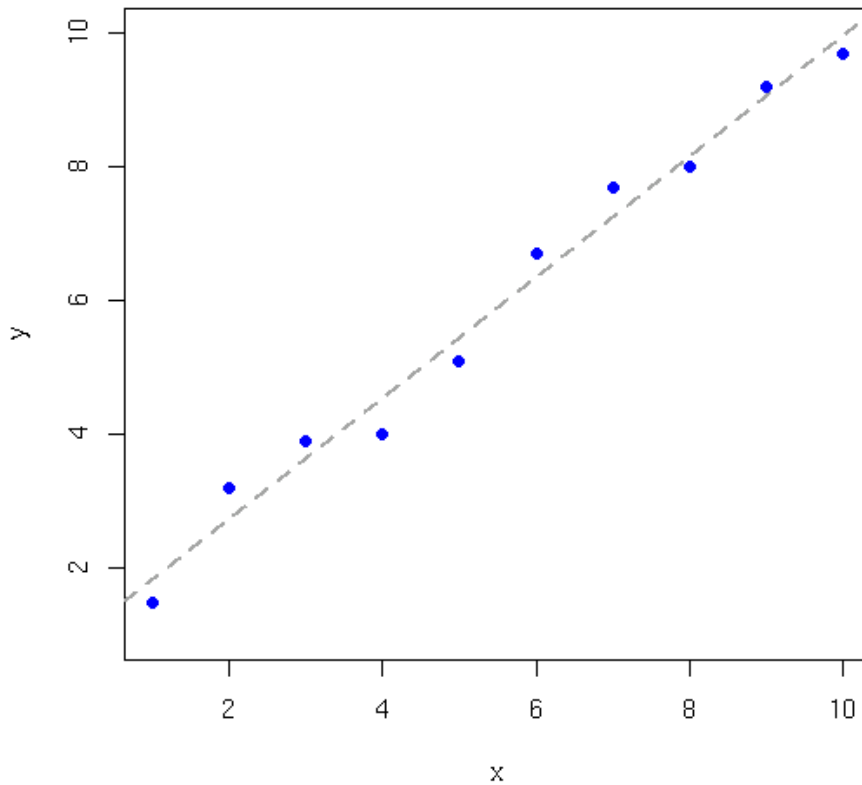
$$\varepsilon_j = C_j - f(\xi, t_j)$$

Μπορούμε να ελαχιστοποιήσουμε την συνάρτηση

$$F(\xi) = \sum_{j=1}^n (C_j - f(\xi, t_j))^2$$

Υπολογίζουμε τις **τιμές των ξ** για τις οποίες η **$F(\xi)$** παίρνει ελάχιστη τιμή
Επίσης υπολογίζουμε το **τυπικό σφάλμα των ξ** ,
δηλαδή την αβεβαιότητα με την οποία εκτιμήθηκαν τα ξ .

Ονομάζεται **Ordinary Least Squares (OLS)**



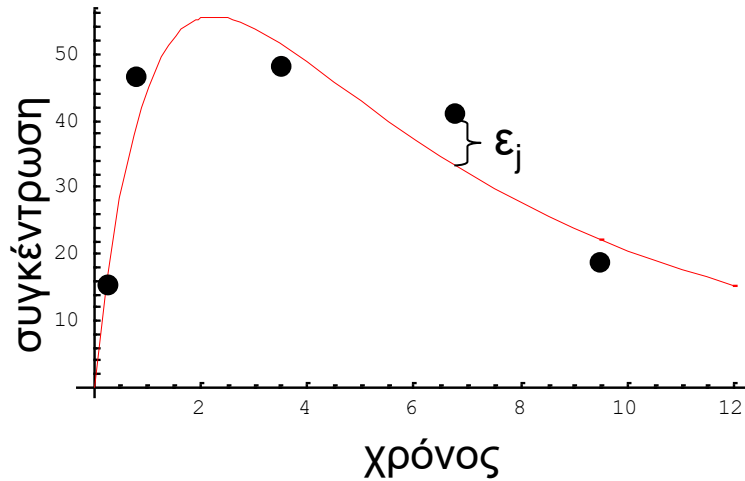
- Επιλογή ΦΚ μοντέλου
- Εκτίμηση παραμέτρων

Το γραμμικό μοντέλο είναι ένα: $y = a x + b$

Υπάρχουν άπειρα μη γραμμικά μοντέλα: $y = f(\xi, x)$

Ποια είναι η συνάρτηση f ;

Μέγιστη πιθανοφάνεια



$$\hat{Y}_j = f(\xi, t_j)$$

$$\varepsilon_j = Y_j - \hat{Y}_j$$

$$\varepsilon_j \sim N(0, \sigma) \quad \text{ή} \quad Y_j \sim N(\hat{Y}_j, \sigma)$$

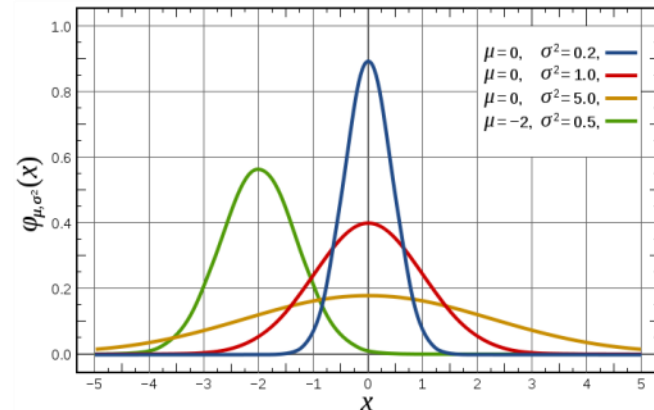
Υπόθεση:

Η πιθανότητα το ε_j (ή το Y_j) να παίρνει μια συγκεκριμένη τιμή ακολουθεί κανονική κατανομή με μέση τιμή 0 (ή \hat{Y}_j) και τυπική απόκλιση σ

Συνάρτηση πυκνότητας πιθανότητας
κανονικής κατανομής

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$f(x)$ είναι η πιθανότητα να παίρνει το X τιμή από x έως $x+dx$



Η πιθανότητα για ένα πειραματικό σημείο να έχει τιμή Y_j ονομάζεται πιθανοφάνεια (likelihood)

$$L(Y_j | \xi, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(Y_j - \hat{Y}_j(\xi))^2}{2\sigma^2}}$$

Για πολλά σημεία: Αν οι πιθανοφάνειες των σημείων είναι ανεξάρτητες μεταξύ τους η πιθανοφάνεια για όλα τα σημεία είναι το γινόμενο τους

$$L(Y | \xi, \sigma) = L(Y_1 | \xi, \sigma_1) \times L(Y_2 | \xi, \sigma_2) \times \dots \times L(Y_n | \xi, \sigma_n) = \prod_{j=1}^n L(Y_j | \xi, \sigma_j)$$

$$L(Y | \xi, \sigma) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(Y_j - \hat{Y}_j)^2}{2\sigma_j^2}}$$

Μπορούμε να υλοποιήσουμε μια αντικειμενική συνάρτηση που πρέπει να ελαχιστοποιείται

$$L(Y | \xi, \sigma) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(Y_j - \hat{Y}_j)^2}{2\sigma_j^2}}$$
$$-2LL(Y | \xi, \sigma) = nLn(2\pi) + \sum_{j=1}^n \left[\frac{(Y_j - \hat{Y}_j)^2}{\sigma_j^2} + Ln(\sigma_j^2) \right]$$

$$F(\xi, \sigma) = \sum_{j=1}^n \left[\frac{(Y_j - \hat{Y}_j)^2}{\sigma_j^2} + Ln(\sigma_j^2) \right]$$

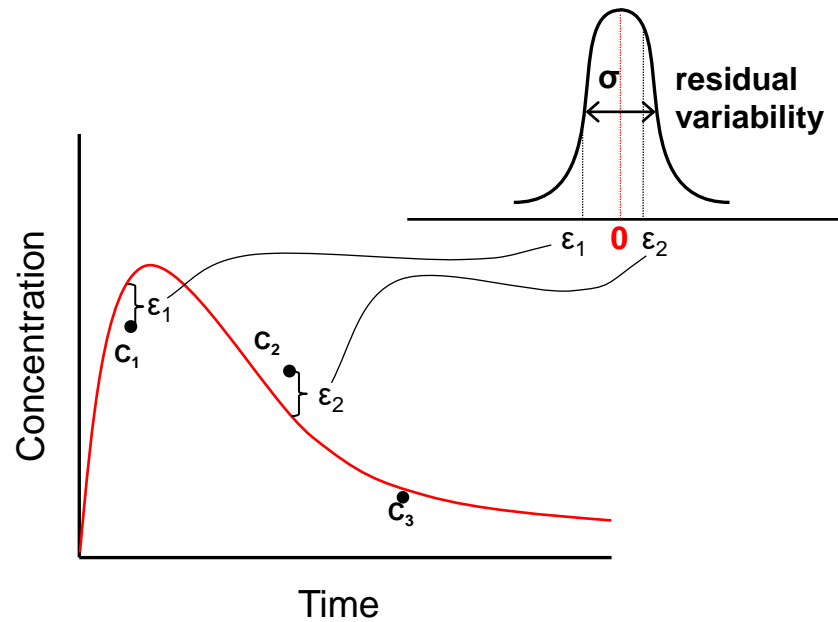
Αυτή η μέθοδος ονομάζεται **Extended Least Squares (ELS)** Τα ξ και σ εκτιμώνται ταυτόχρονα

Στην ειδική περίπτωση του ομοσκεδαστικού σφάλματος όπου $\sigma_j = \sigma$
Η αντικειμενική συνάρτηση είναι **OLS**

$$F(\xi) = \sum_{j=1}^n (Y_j - \hat{Y}_j)^2 \quad \text{και} \quad \sigma^2 = \frac{\sum_{j=1}^n (Y_j - \hat{Y}_j)^2}{n}$$

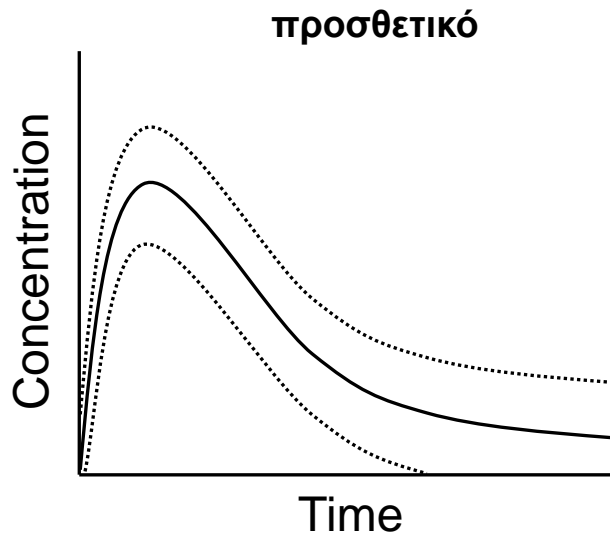
Μοντέλο σφάλματος

$$C_j = f(\xi, t_j) + \varepsilon_j$$

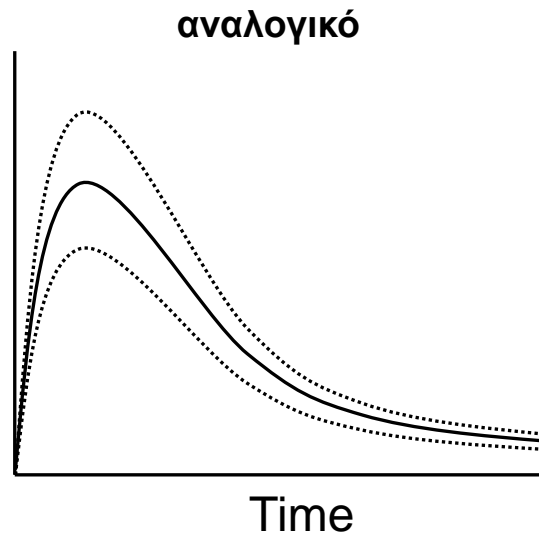


Απλό προσθετικό (ομοσκεδαστικό) σφάλμα $\varepsilon \sim N(0, \sigma)$

Όταν δεν είναι ομοσκεδαστικό είναι ετεροσκεδαστικό σφάλμα



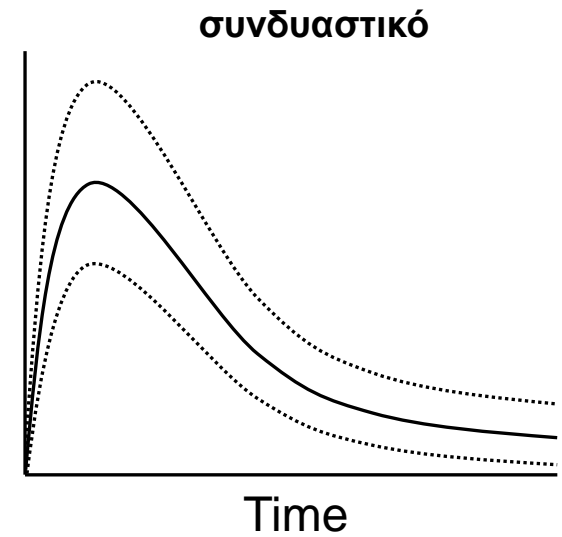
$$C = f + \varepsilon$$



$$C = f + f \cdot \varepsilon$$

εναλλακτικά

$$C = f \cdot e^{\varepsilon}$$

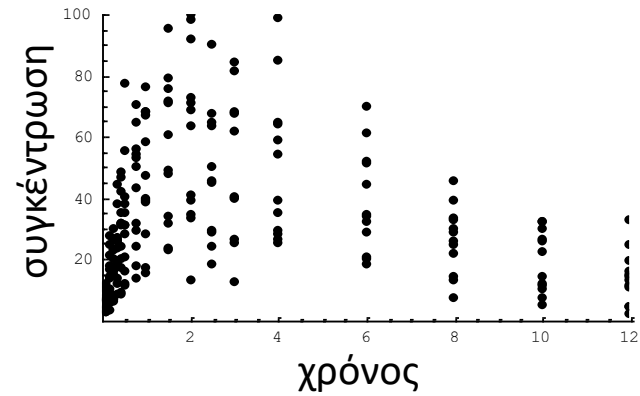
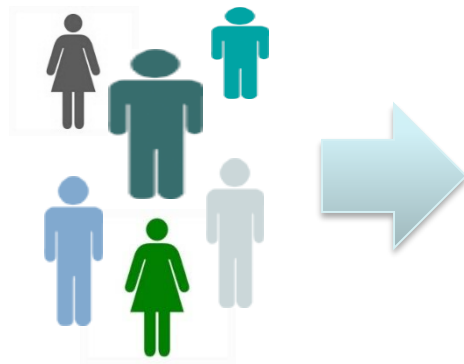


$$C = f + f \cdot \varepsilon_1 + \varepsilon_2$$

- Επιλογή ΦΚ μοντέλου
- Επιλογή μοντέλου σφάλματος
- Εκτίμηση παραμέτρων

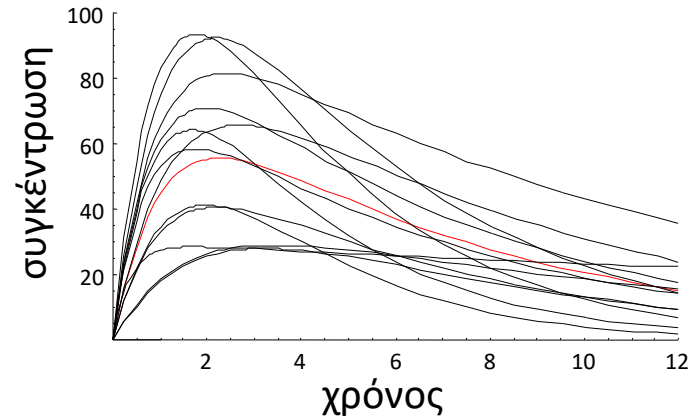
Πληθυσμιακή φαρμακοκινητική

Δι-ατομική μεταβλητότητα ΦΚ παραμέτρων



Παραμετρική πληθυσμιακή φαρμακοκινητική

δι-ατομική μεταβλητότητα



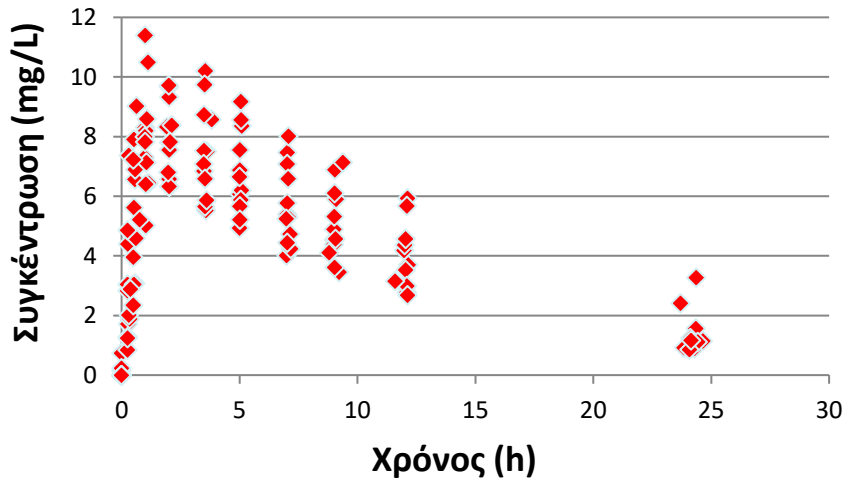
$$C_j = f(\xi, t_j) + \varepsilon_j$$



$$C_{ij} = f(\xi_i, t_j) + \varepsilon_{ij}$$

Μέθοδος 2 σταδίων

1^ο στάδιο: ατομικές προσαρμογές



ID	KA	V	CL
1	1.77	27.33	1.64
2	2.12	33.33	3.00
3	2.46	34.63	2.77
4	1.25	32.38	2.59
5	1.52	28.88	2.31
6	1.22	41.78	3.76
7	0.75	34.11	3.07
8	1.43	34.78	3.13
9	6.75	34.50	2.76
10	0.71	26.71	1.87
11	3.62	35.70	3.57
12	0.97	25.44	2.29

2^ο στάδιο: περιγραφική στατιστική

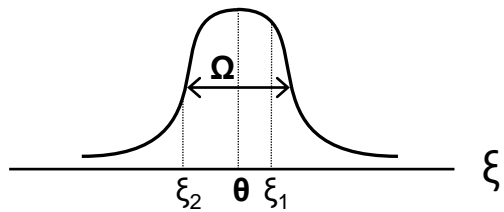
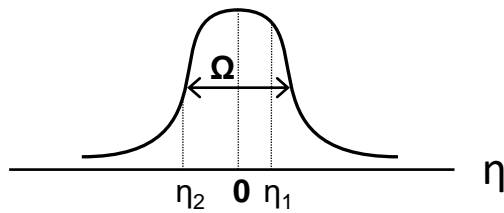
	KA	V	CL
Μέσος	2.0475	32.46	2.73
Τυπ. αποκλ (CV)	82%	14%	23%

Μειονέκτημα: Πρέπει να έχουμε αρκετά δεδομένα από τον κάθε ασθενή

Μοντέλα μικτών επιδράσεων (Mixed effects models)

κατανομή παραμέτρου ξ

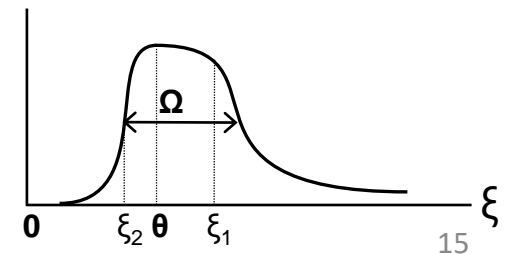
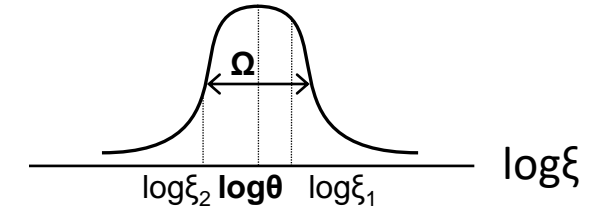
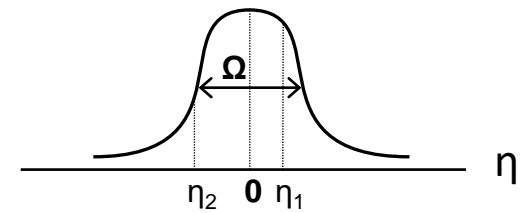
normal
 $\xi_i = \theta + \eta_i$



μέση τιμή: θ
πληθυσμιακή τιμή
διακύμανση: Ω
δι-ατομική μεταβλητότητα

$$\eta_i \sim N(0, \Omega)$$

log-normal
 $\xi_i = \theta e^{\eta_i} = \theta \exp(\eta_i)$
 $\log \xi_i = \log \theta + \eta_i$



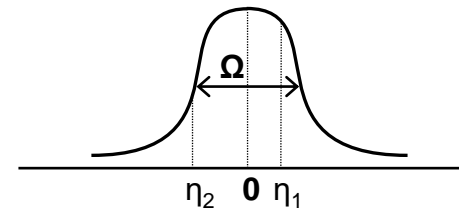
Ιεραρχικό μοντέλο με 2 (ή περισσότερα) επίπεδα μεταβλητότητας

φαρμακοκινητικό μοντέλο

$$C_{ij} = f(\theta e^{\eta_i}, t_{ij}) + \varepsilon_{ij}$$

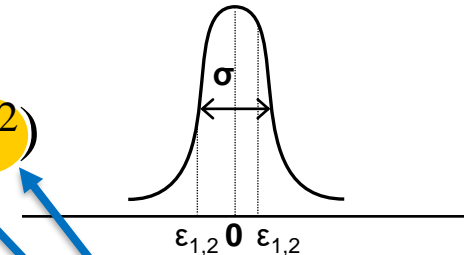
δι-ατομική μεταβλητότητα

$$\eta_i \sim N(0, \Omega)$$

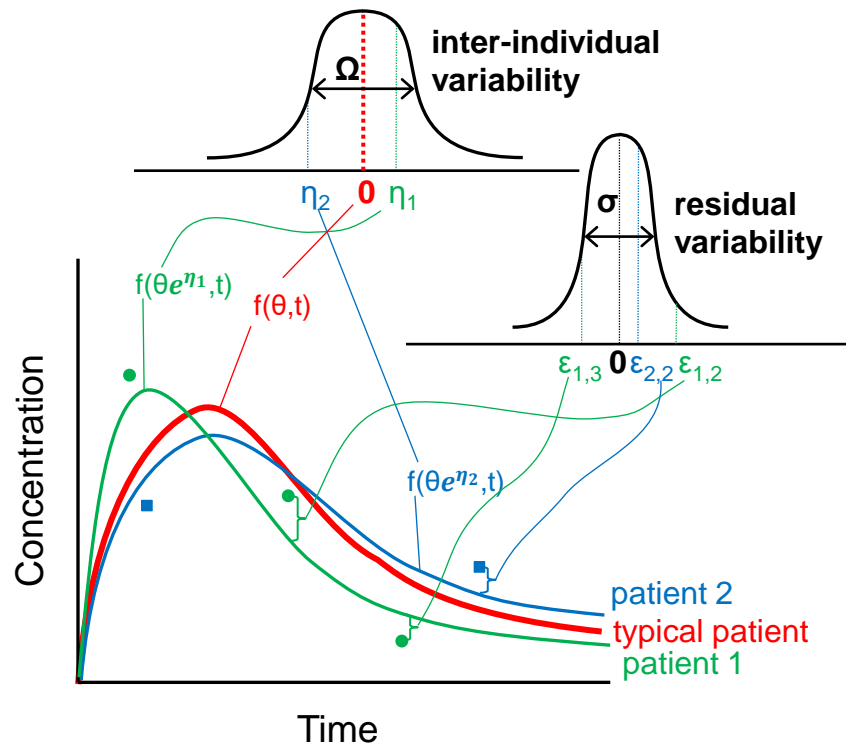


υπόλοιπη μεταβλητότητα

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$



Ταυτόχρονη προσαρμογή του ΦΚ στα δεδομένα όλων των ασθενών με σκοπό την εκτίμηση των **fixed effects: θ** , και **random effects: Ω, σ** καθώς και τα τυπικά σφάλματα αυτών



Συσχετίσεις μεταξύ των παραμέτρων

Μπορεί να υπάρχουν συσχετίσεις μεταξύ των παραμέτρων
Τότε η διατομική μεταβλητότητα είναι ένας πίνακας variance – covariance

Η μη διαγώνιοι όροι εκφράζουν τις συν-διακυμάνσεις μεταξύ των παραμέτρων

Συχνά εκφράζονται και ως συσχετίσεις.

Εν γένει αναμένεται να υπάρχουν συσχετίσεις όμως συχνά τα δεδομένα δεν έχουν αρκετή πληροφορία ώστε αυτές να εκτιμηθούν.

Π.χ. Συνήθεις παράμετροι που συσχετίζονται στην ΦΚ είναι η κάθαρση και ο όγκος κατανομής

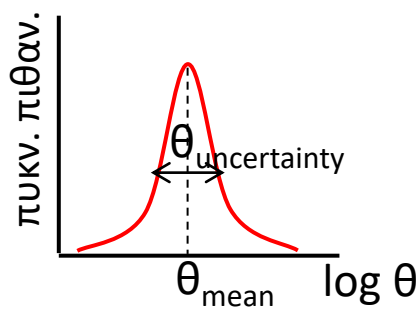
Στα πλαίσια της θεωρίας της μέγιστης πιθανοφάνειας
Κατασκευάζουμε μια συνάρτηση πιθανοφάνειας που δεν εκφράζεται
αναλυτικά και περιγράφει με ένα ιεραρχικό μοντέλο τα παρακάτω:

Τις μέσες τιμές θ
Την δι-ατομική μεταβλητότητα Ω
Την ενδο-ατομική μεταβλητότητα σ

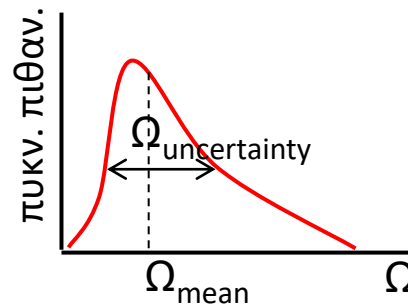
Οι παράμετροι αυτές εκτιμώνται όλες μαζί με το μοντέλο μικτών
επιδράσεων και υπολογίζονται:

οι μέσες τιμές των θ , Ω και σ

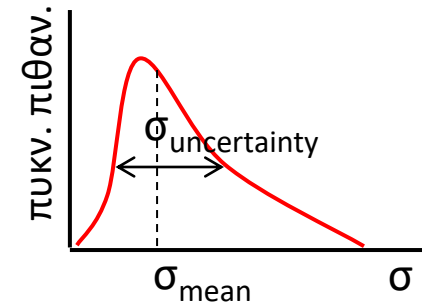
και οι αβεβαιότητες (standard errors) αυτών που ακολουθούν κατανομές



Log-normal



Inverse gamma (univariate)
Inverse Wishart (multivariate)



Inverse gamma

Διάκριση Μεταβλητότητας και Αβεβαιότητας

Εκφράζονται και οι 2 από στατιστικές κατανομές,

- έχουν μέση τιμή και διακύμανση
- όμως είναι διαφορετικές έννοιες

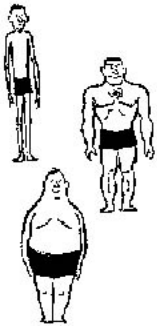
Δι-ατομική μεταβλητότητα, Μέση τιμή: θ , διακύμανση: Ω	Χαρακτηριστικό του πληθυσμού	Πιο πολλά άτομα -> σταθερό
Υπόλοιπη μεταβλητότητα, Μέση τιμή: 0, διακύμανση: σ^2	Χαρακτηριστικό του πειράματος	Πιο πολλά δείγματα -> σταθερό
Αβεβαιότητα του θ και του Ω Μέση τιμή: θ ή Ω διακύμανση: τυπικό σφάλμα θ ή Ω	Εξαρτάται από τον αριθμό των ατόμων (και των δειγμάτων)	Πιο πολλά άτομα -> μειώνεται
Αβεβαιότητα του σ Μέση τιμή: 0, διακύμανση: τυπικό σφάλμα σ	Εξαρτάται από τον αριθμό των δειγμάτων	Πιο πολλά δείγματα -> μειώνεται

Περιγραφικές συμμεταβλητές (covariates)

- Στην πραγματικότητα ΟΛΗ η μεταβλητότητα που παρατηρείται οφείλεται σε διάφορους παράγοντες που καθιστούν το κάθε άτομο διαφορετικό
- Οι περισσότεροι παράγοντες είναι άγνωστοι, έτσι ποσοτικοποιούνται με στοχαστικό τρόπο (δι-ατομική μεταβλητότητα, Ω)
- Κάποιοι όμως παράγοντες μπορεί να είναι γνωστοί και εισάγονται στο μοντέλο, όπως **βάρος, γενετική πληροφορία, ηλικία, εργαστηριακές μετρήσεις, κλπ.**
- Η συνολική δι-ατομική μεταβλητότητα της ΦΚ παραμέτρου δεν αλλάζει, απλά αντί να περιγράφεται μόνο από μια στοχαστική μεταβλητή, μοιράζεται μεταξύ των συμμεταβλητών και της στοχαστικής μεταβλητής

Οι συμμεταβλητές που ενσωματώνονται στο μοντέλο και ερμηνεύουν τμήμα της μεταβλητότητας με αιτιοκρατικό τρόπο μπορεί να είναι διαφόρων τύπων, παραδείγματα:

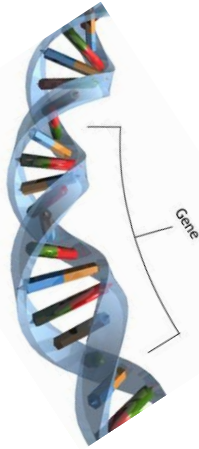
σωματομετρικές



εθνικές



γενετικές



εργαστηριακές



συγχορηγήσεις



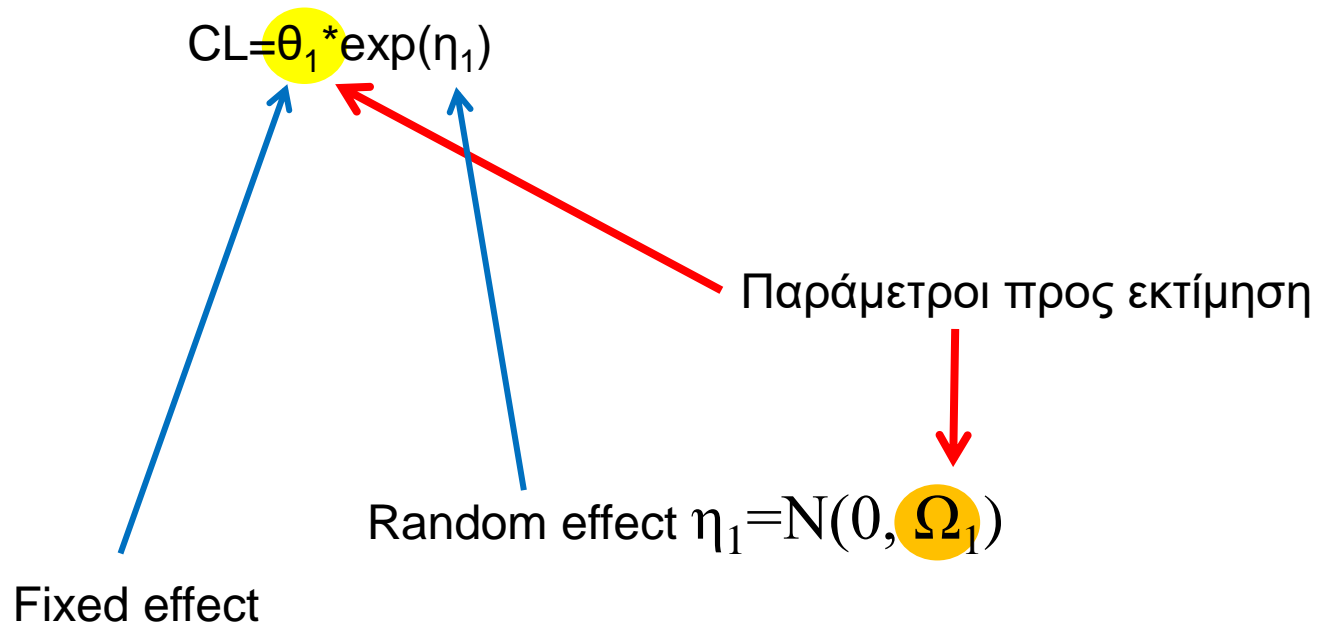
περιβάλλον



Περιγραφικές συμμεταβλητές στην πληθυσμιακή ΦΚ

Στην πληθυσμιακή ΦΚ μπορούμε εύκολα να συμπεριλάβουμε συμμεταβλητές

Χωρίς συμμεταβλητή η ΦΚ παράμετρος CL γράφεται:



Περιγραφικές συμμεταβλητές στην πληθυσμιακή ΦΚ

Στην πληθυσμιακή ΦΚ μπορούμε εύκολα να συμπεριλάβουμε συμμεταβλητές

Χωρίς συμμεταβλητή η ΦΚ παράμετρος CL γραφεται:

$$CL = \theta_1 * \exp(\eta_1)$$

Περιγραφικές συµµεταβλητές στην πληθυσµιακή ΦΚ

Στην πληθυσµιακή ΦΚ µπορούµε εύκολα να συµπεριλάβουµε συµµεταβλητές

Χωρίς συµµεταβλητή η ΦΚ παράµετρος CL γραφεται:

$$CL = \theta_1 * \exp(\eta_1)$$

Με συµµεταβλητή το βάρος:

$$CL = (\theta_1 + \theta_2 * WT) * \exp(\eta_1)$$

Fixed effects

Συµµεταβλητή (βάρος)

Παράµετροι προς εκτίµηση

Random effect $\eta_1 = N(0, \Omega_1)$

Συνήθως χρησιμοποιούμε γραμμικά μοντέλα

Συνήθως γράφουμε τον covariate centred ως προς μια τυπική τιμή του, \overline{cov} έτσι ώστε για τον τυπικό ασθενή με $cov_1 = \overline{cov_1}$, να είναι $Par = \theta_1$

Συνήθη μοντέλα συμμεταβλητών

Γραμμικό (continuous)

$$Par = (\theta_1 + \theta_2(cov - \overline{cov})) \exp \eta_1$$

Power (continuous)

$$Par = \theta_1 \left(\frac{cov}{\overline{cov}} \right)^{\theta_2} \exp \eta_1$$

$$\log P = \log \theta_1 + \theta_2(\log cov - \log \overline{cov}) + \eta_1$$

Categorical covariate

$$Par = (\theta_1 + \theta_2 * cov) \exp \eta_1$$

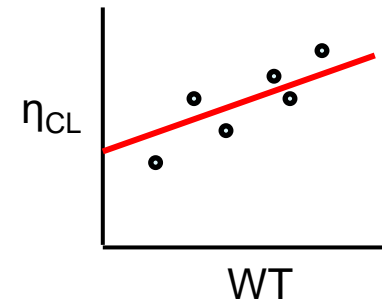
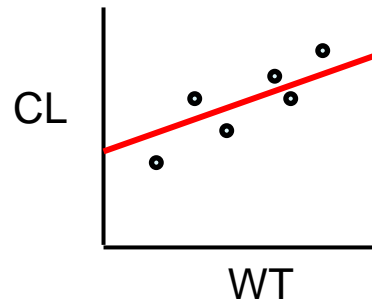
$$Par = \theta_1 * \theta_2^{cov} \exp \eta_1$$

Οι συμμεταβλητές πρέπει να εντοπιστούν

Αρχικά
μοντελοποιούμε
χωρίς συμμεταβλητές

$$CL = \theta_1 * \exp(\eta_1)$$

Αναζητούμε συσχέτιση
π.χ. του WT με τις
ατομικές τιμές CL



Θα πούμε στη συνέχεια
που βρίσκουμε τα CL

$$CL = (\theta_1 + \theta_2 * (WT - WT_{mean})) * \exp(\eta_1)$$

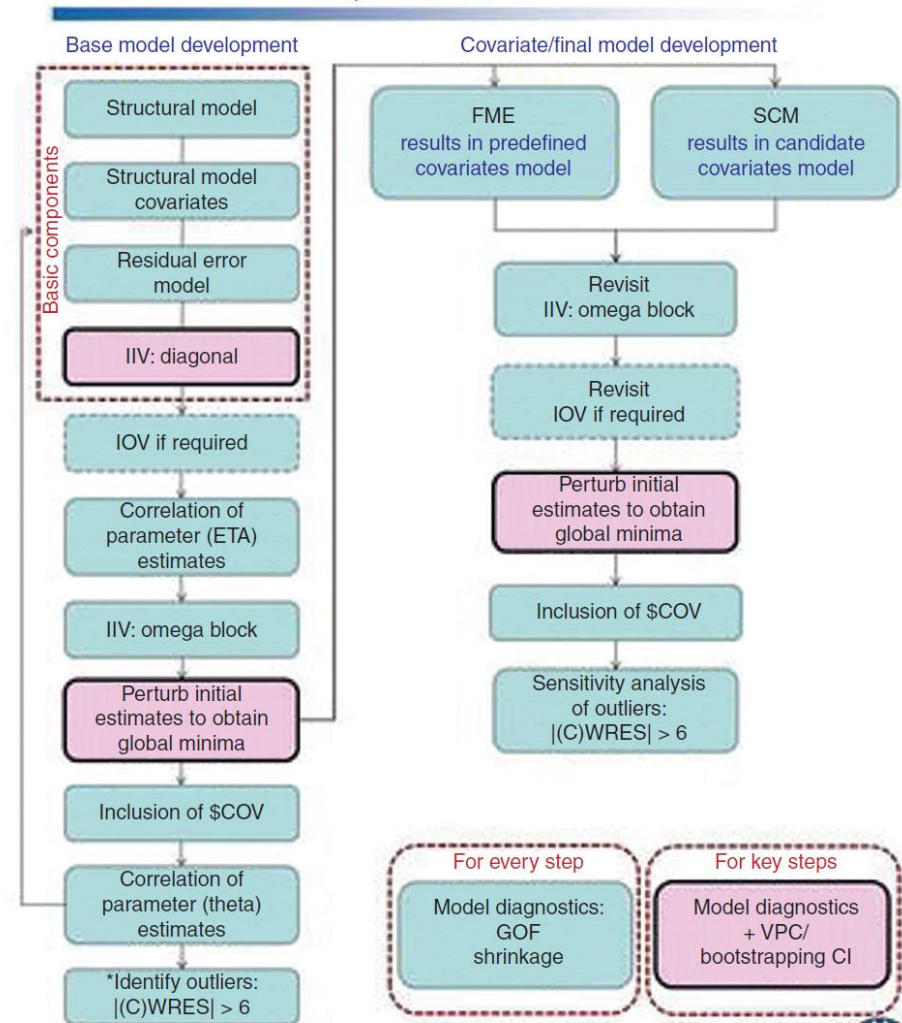
Θα πούμε στη συνέχεια για
στρατηγικές εντοπισμού

- Επιλογή ΦΚ μοντέλου
- Επιλογή μοντέλου σφάλματος
- Επιλογή πληθυσμιακού μοντέλου**
- Εκτίμηση παραμέτρων

Ανάπτυξη πληθυσμιακού μοντέλου

Pop PK workflow

- ✓ Επιλογή ΦΚ μοντέλου
- ✓ Επιλογή μοντέλου σφάλματος
- ✓ Επιλογή πληθυσμιακού μοντέλου (με συμμεταβλητές)
- ✓ Εκτίμηση παραμέτρων
- ✓ Επικύρωση του μοντέλου



* If not proceeding to covariate model building, perform sensitivity analysis.



Αξιολόγηση μοντέλου

- ✓ Επιτυχής ολοκλήρωση του αλγόριθμου
- ✓ Λογικές τιμες των παραμέτρων (με φυσική σημασία)
- ✓ Λογικά (μικρα) τυπικά σφάλματα
- ✓ Διαγνωστικά γραφήματα (goodness of fit)
- ✓ Κριτήρια επιλογής μοντέλου
- ✓ Επικύρωση μοντέλου
 - Visual predictive check
 - Bootstrap

Συντελεστής προσδιορισμού, R^2
Coefficient of determination

$$R^2 = 1 - \frac{\sum_{j=1}^n (Y_j - \hat{Y}_j)^2}{\sum_{j=1}^n (Y_j - \bar{Y})^2}$$

Όπου \bar{Y} είναι η μέση τιμή των πειραματικών δεδομένων

Η σημασία του R^2 είναι το ποσοστό της μεταβλητότητας των δεδομένων που ερμηνεύεται από το μοντέλο

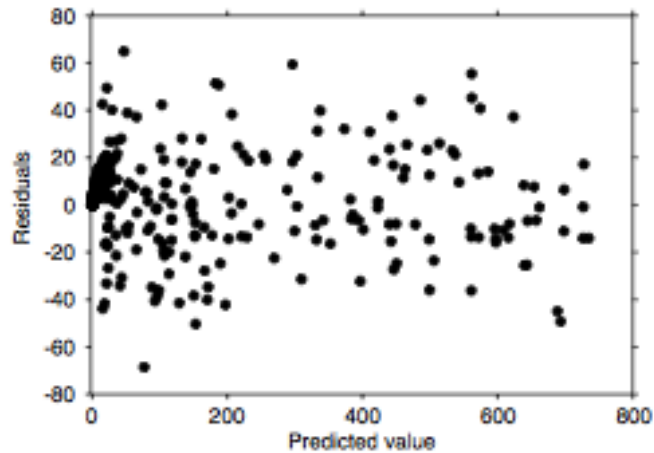
Ιδανικά $R^2 = 1$

Στη χειρότερη $R^2 = 0$

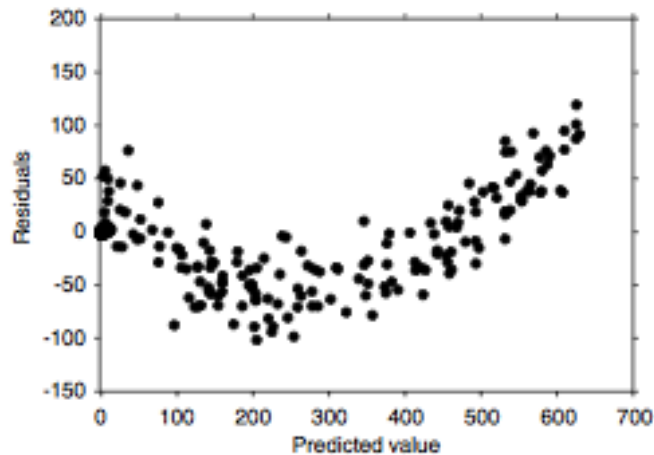
Διαγνωστικά γραφήματα

Τα διαγνωστικά γραφικά είναι πιο πληροφοριακά από τα κριτήρια με έναν αριθμό

Residuals vs Predicted

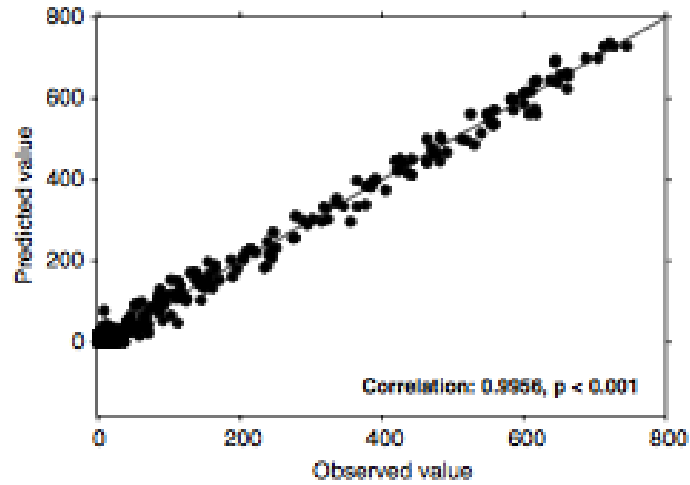


καλό

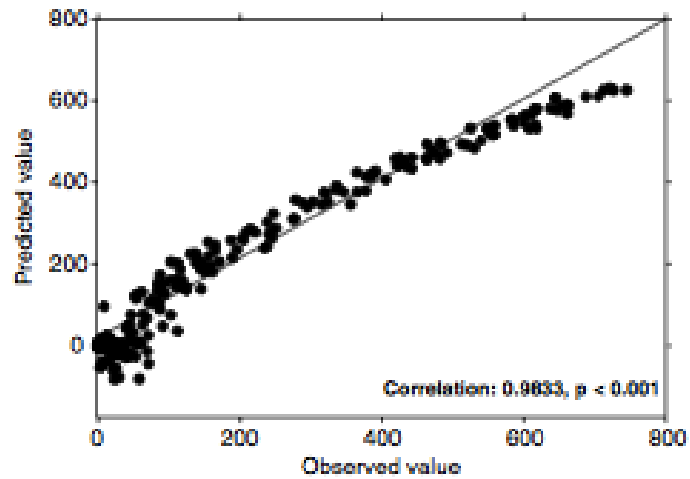


προβληματικό

Observed vs Predicted



καλό



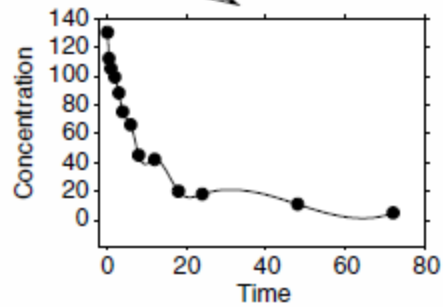
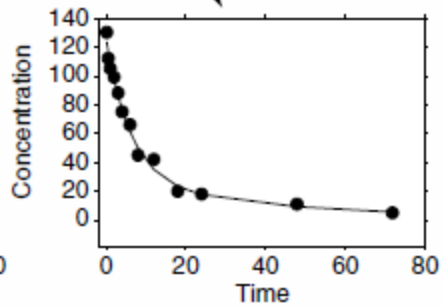
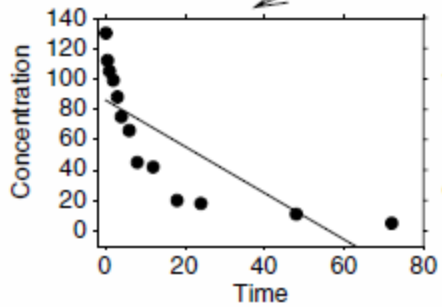
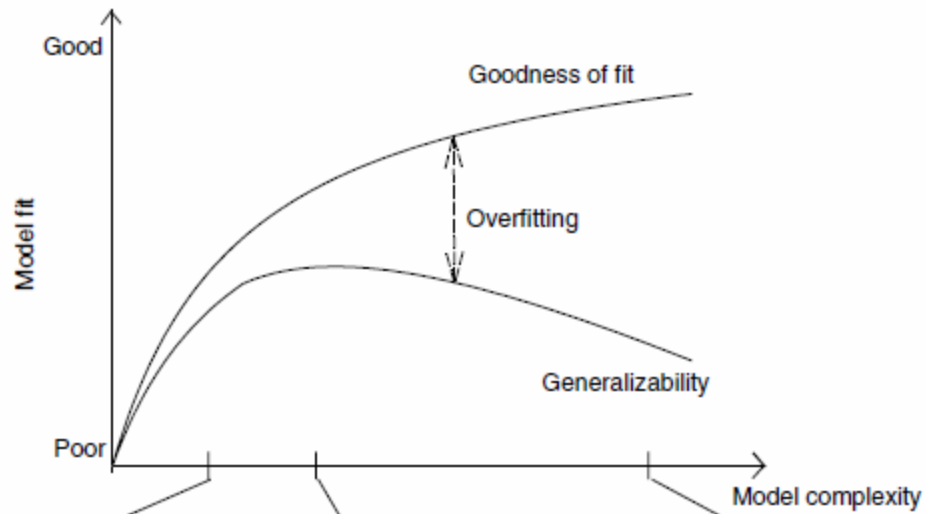
προβληματικό

Επιλογή μοντέλου

Η επιλογή του μοντέλου είναι η πιο σημαντική απόφαση
Τα αλλά τα κάνει ο υπολογιστής...

Ένα πιο πολύπλοκο μοντέλο θα περνάει πιο κοντά από τα δεδομένα....
... αλλά θα έχει πιο πολλές παραμέτρους που θα εκτμηθούν χειρότερα





Κριτήρια επιλογής μοντέλου

Όταν υπάρχουν πολλά υποψήφια μοντέλα μπορούμε να χρησιμοποιήσουμε κριτήρια επιλογής του καλύτερου

Κριτήρια πληροφορίας

Κριτήρια που βασίζονται στη θεωρία πληροφορίας

Πχ Akaike Information Criterion (AIC),
Bayesian Information Criterion (BIC)

Στατιστικά κριτήρια

Κριτήρια που βασίζονται σε στατιστικά τεστ (hypothesis tests)
Πχ Likelihood ratio test

Akaike Information Criterion (AIC)

- Βασιζόμενο στην πληροφορία Kullback – Leibler διατυπώνεται το:

$$AIC = -2LL + 2p$$

LL: Log-likelihood

p: Αριθμός παραμέτρων

- Επιλέγουμε το μοντέλο με τη μικρότερη τιμή AIC
- Το AIC αποτελεί συμβιβασμό ανάμεσα σε προσαρμογή και αριθμό παραμέτρων

Bayesian Information Criterion (BIC ή Schwarz criterion)

$$BIC = -2LL + pLn(n)$$

- Όπως και με το AIC επιλέγουμε το μοντέλο με τη μικρότερη τιμή BIC
- Το BIC είναι πιο συντηρητικό από το AIC και συχνά ευνοεί το απλούστερο μοντέλο

Χαρακτηριστικά AIC και BIC

- Τα AIC και BIC, για ίδιο αριθμό παραμέτρων προκρίνουν το μοντέλο με την καλύτερη προσαρμογή
- Το p αναφέρεται σε όλες τις εκτιμώμενες παραμέτρους (πχ σ) όχι μόνο στις παραμέτρους του δομικού μοντέλου (ξ).
- Τα AIC και BIC, μπορούν να χρησιμοποιηθούν και σε μοντέλα που δεν είναι φωλιασμένα (nested, θα δούμε τι είναι αυτά)
- Τα AIC και BIC, πάντα χρησιμοποιούνται για τα ίδια δεδομένα

Φωλιασμένα μοντέλα (nested models)

Δηλαδή μοντέλα που το ένα είναι ειδική περίπτωση του άλλου

Π.χ. $Y = b * X + a$
 $Y = b * X$

Επίσης τα διαμερισματικά μοντέλα στην ΦΚ συνήθως είναι nested
Όμως το δομικό μοντέλο το επιλέγουμε με κριτήρια goodness of fit.

Εκεί που χρησιμεύουν περισσότερο τα στατιστικά τεστ είναι στην επιλογή
στατιστικά σημαντικών covariates

Τεστ των λόγων πιθανοφάνειας Likelihood ratio test

Έστω 2 μοντέλα φωλιασμένα, ένα με αριθμό παραμέτρων f και ένα με r , έτσι ώστε $f > r$. Το likelihood ratio (λ) ακολουθεί κατανομή χ^2 :

$$\lambda = \frac{L_r}{L_f}$$

Το “r” (reduced) είναι η μηδενική υπόθεση και το “f” (full) η εναλλακτική

Το “f” (full) μοντέλο είναι καλύτερο με επίπεδο εμπιστοσύνης α (πχ $\alpha=0.05$), δηλαδή απορρίπτεται η μηδενική υπόθεση, όταν

$$-2\text{Ln}(\lambda) \geq \chi_{(f-r),\alpha}^2$$

$$2(LL_f - LL_r) \geq \chi_{(f-r),\alpha}^2$$

- Το -2LL μας το δίνει το πρόγραμμα για τα 2 μοντέλα
- Την τιμή του χ^2 την διαβάζουμε από πίνακες
- Για 1 παράμετρο και $\alpha=0.05$ η διαφορά των 2 τιμών πρέπει να είναι >3.84

Degrees of freedom (<i>df</i>)	χ^2 value [9]											
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83	
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82	
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27	
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47	
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52	
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46	
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32	
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12	
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88	
10	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59	
P value (Probability)	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001	
	Nonsignificant								Significant			

Το LRT εφαρμόζεται μόνο σε nested μοντέλα

Στρατηγικές εντοπισμού covariates (Stepwise covariate modeling)

Forward detection – Backward deletion με LRT

- Οι covariates προστίθενται ένας – ένας έτσι ώστε:
(Obj func val full model) – (Obj func val reduced model) < -3.85 μονάδες (0.05 επίπεδο εμπιστοσύνης).
- Στο τέλος οι covariates αφαιρούνται ένας-ένας και παραμένουν στο μοντέλο αν:
(Obj func val full model) – (Obj func val reduced model) > 10.83 μονάδες (0.001 επίπεδο εμπιστοσύνης).

Εκτος από το LRT κάθε covariate ελέγχεται για εξής:

- Έχει μία λογική φυσική σημασία;
- Το αντίστοιχο random effect (IIV) ελαττώνεται; δηλαδή ερμηνεύεται από τον covariate ένα τμήμα της στοχαστικής μεταβλητότητας;
- Έχει κλινικά σημαντική επίδραση εντός του εύρους των τιμών του covariate;

$$\text{π.χ. } \frac{CL(\max \text{ cov}) - CL(\min \text{ cov})}{CL(\max \text{ cov})} \times 100 > 10\%$$

- Προσέχουμε δευτερογενείς συσχετίσεις π.χ το βάρος σχετίζεται με το ύψος ή την επιφάνεια σώματος, άρα προσέχουμε πως θα μπουν και τα 2.

Στρατηγικές εντοπισμού covariates (full model estimation)

Wald test

Αποτελεί προσέγγιση του LRT και υπολογίζεται από το full model μέσω των standard errors των παραμέτρων που αφορούν τους covariates.

Αναφέρει P values για το αν η παράμετρος του κάθε covariate είναι σημαντικά διαφορετική από το 0.

Είναι εύχρηστο και γρήγορο (αν δουλέψει) αλλά δεν είναι πάντα εύκολο να γίνει αξιόπιστα εκτίμηση των παραμέτρων του full model και με standard errors.

Άλλες μέθοδοι

Π.χ. **LASSO** (least absolute shrinkage and selection operator)

Μια μέθοδος παλινδρόμησης που εισάγοντας κατάλληλο constrain, κάποιες από τις εκτιμώμενες παραμέτρους μηδενίζονται, οπότε ξεκινώντας από ένα full model, μπορεί να καταλήξει σε ένα reduced model κρατώντας μόνο τις παραμέτρους που δεν μηδενίζονται.

Internal Validation

Η **εσωτερική** επικύρωση του μοντέλου είναι να διαπιστώσουμε ότι το μοντέλο περιγράφει ικανοποιητικά τα δεδομένα από τα οποία προήλθε

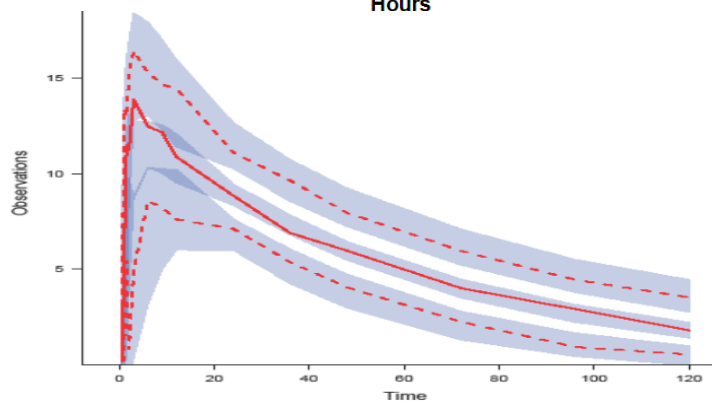
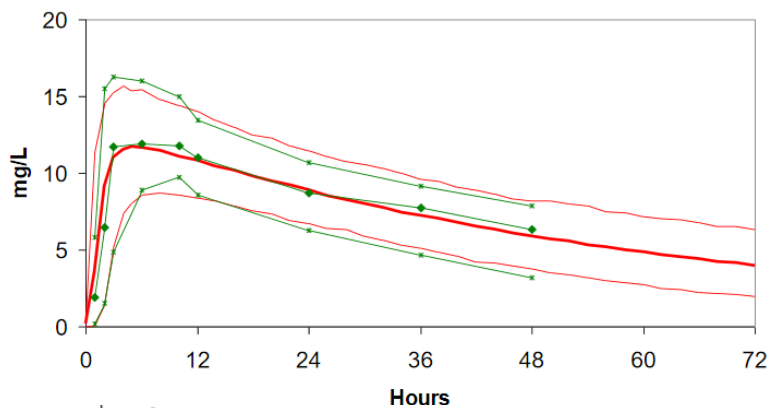
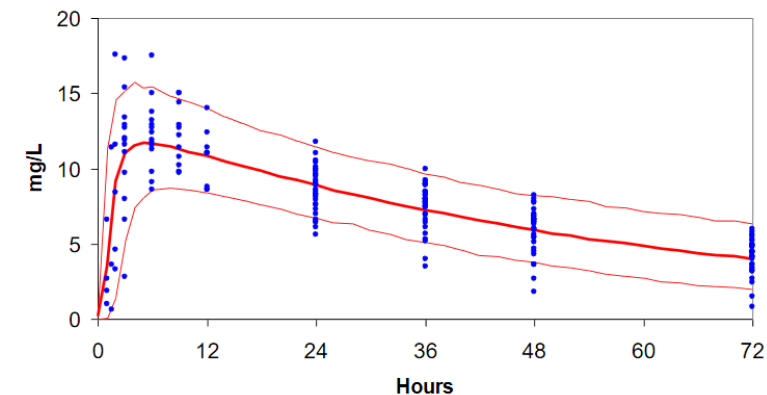
Visual Predictive Check

Numerical Predictive Check

Bootstrapping

Η **εξωτερική** είναι να περιγράφει ικανοποιητικά νέα δεδομένα που **δεν** χρησιμοποιήθηκαν για την ανάπτυξή του.

Visual Predictive Check (VPC)



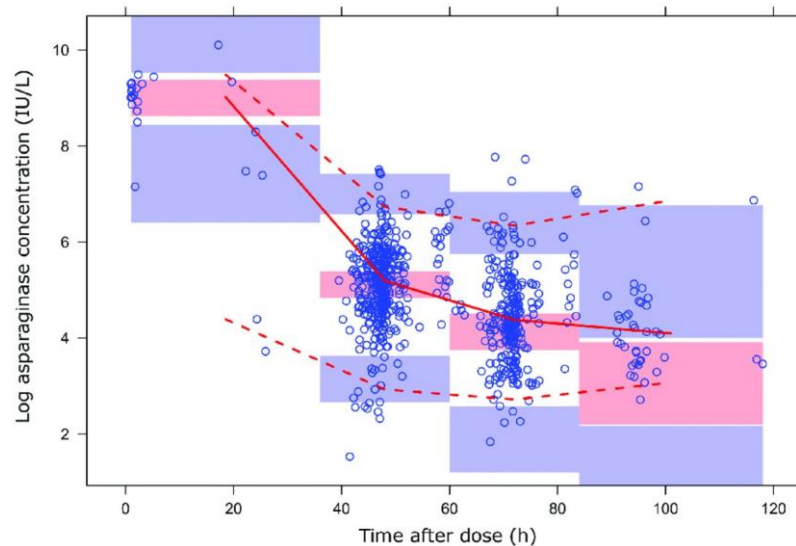
- Προσομοίωση πολλαπλών (1000) σετ δεδομένων ίδιας δομής με τον αρχικό. Αριθμός ασθενών , δόσεις, Χρόνοι, covariates κλπ.
- Κάθε σετ προσομοιώνεται με όλα τα random effects με τυχαίες τιμές από τις πληθυσμιακές κατανομές (προστίθεται και το residual variability)
- Οι προσομοιώσεις συγκρίνονται με τα δεδομένα ως εξής:
- Υπολογίζονται για κάθε χρονική στιγμή 3 percentiles (5,50,95%) των δεδομένων αλλά και των προσομοιώσεων (PI)
- Υπολογίζονται για κάθε χρονική στιγμή τα 95% διαστήματα εμπιστοσύνης των percentiles των προσομοιώσεων.
- Ελέγχεται αν τα percentiles των δεδομένων βρίσκονται εντός των ΔΕ των αντιστοίχων percentiles των προσομοιώσεων.

Binning

Τα VPC απεικονίζουν τα PI του μοντέλου μαζί με τα δεδομένα συναρτήσει μιας ανεξάρτητης μεταβλητής, συνήθως του χρόνου.

Όμως η ανεξάρτητη μεταβλητή μπορεί να είναι άλλη (πχ δόση).

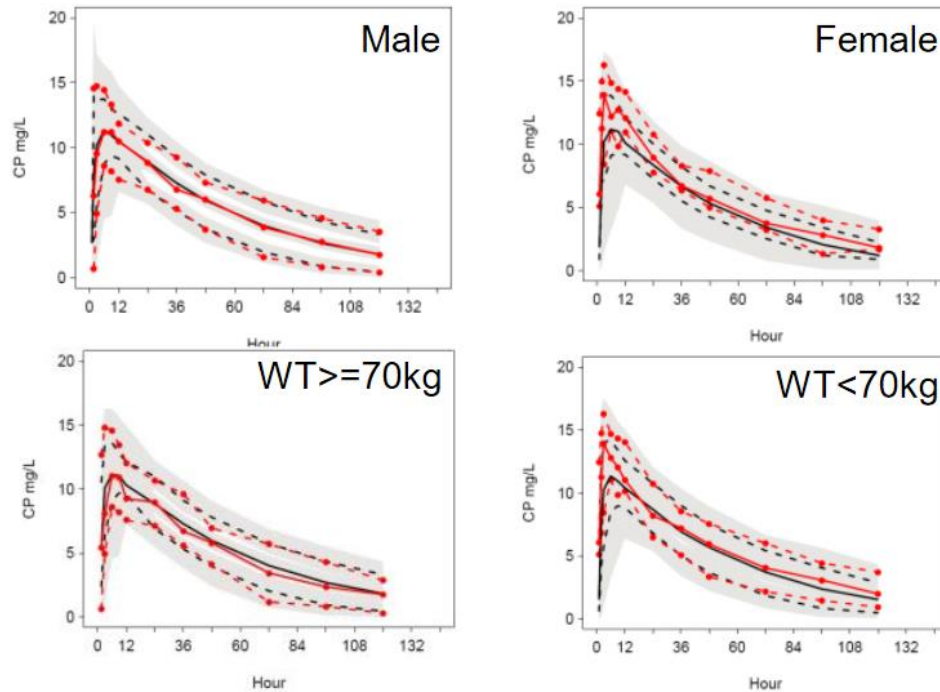
Στα VPC τόσο ως προς χρόνο (ή ως προς άλλη μεταβλητή), οι τιμές του χρόνου μπορεί να είναι σκόρπιες και όχι κοινές για όλους τους ασθενείς. Για να μπορούν να υπολογιστούν τα percentiles κάνουμε ομαδοποίηση των τιμών **binning**



Stratification

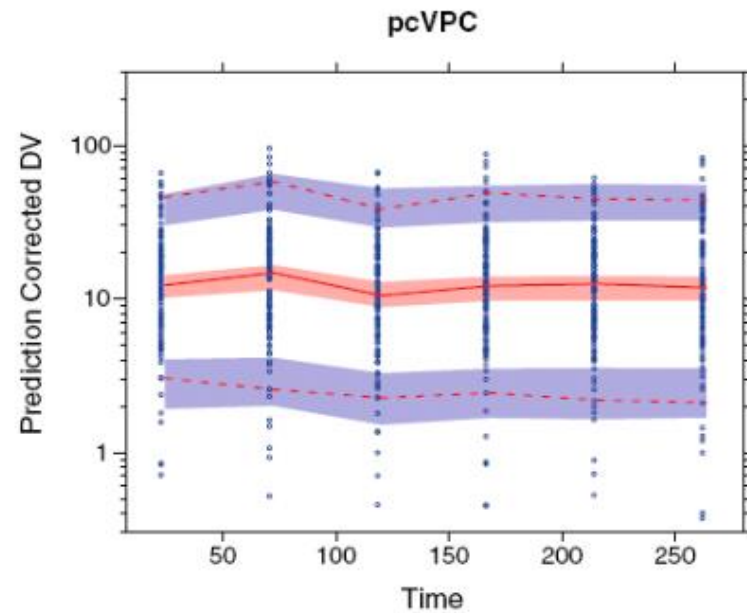
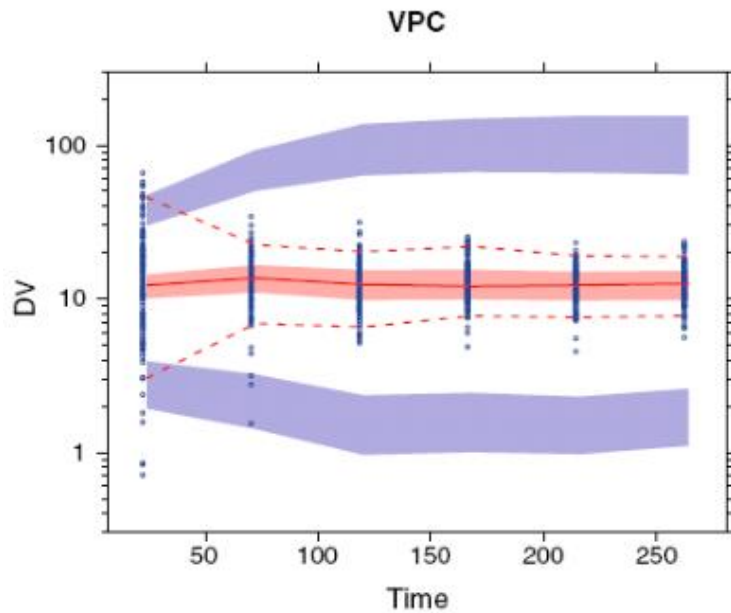
Μία δεύτερη ανεξάρτητη μεταβλητή την απεικονίζουμε με πολλαπλά VPC, που χωρίζουν στρωματοποιημένα τα δεδομένα μας

Επίσης η επίδραση ενός covariate (ανεξ. μεταβλ.) αν είναι μεγάλη μπορεί να καθιστά τα δεδομένα μας μη συγκρίσιμα οπότε είναι λάθος να τα ομαδοποιήσουμε όλα μαζί και πρέπει να τα χωρίσουμε σε υποομάδες



Prediction corrected VPC

Εκτός από το stratification μια άλλη μέθοδος για να λάβουμε υπ' όψιν σε ένα διάγραμμα, ένα covariate μεγάλη επίδραση (πχ δόση) ή γενικότερα όταν έχουμε πολύ ετερογενή δεδομένα είναι να **κανονικοποιήσουμε** και τα PRED και τα OBS με το median των PRED σε κάθε bin



Bootstrap

Γενικά η μέθοδος bootstrap αναφέρεται σε δημιουργία νέων dataset με τυχαία επαναδειγματοληψία από το αρχικό dataset.

Δημιουργούμε μεγάλο αριθμό (1000) νέων δεδομένων ίδιας δομής με το αρχικό dataset, που το κάθε ένα περιλαμβάνει τα δεδομένα τυχαίων ασθενών, ίδιου αριθμού ασθενών με το αρχικό dataset.

original dataset	bootstrap #1	bootstrap #2	bootstrap #3	bootstrap #4	bootstrap #5	bootstrap #6
1	9	4	1	2	5	2
2	3	9	1	8	1	10
3	9	6	6	4	3	1
4	3	6	8	6	10	8
5	10	10	10	2	2	9
6	4	3	2	7	9	9
7	2	8	6	3	6	1
8	3	8	5	7	10	4
9	7	4	1	7	1	3
10	5	6	4	8	5	9

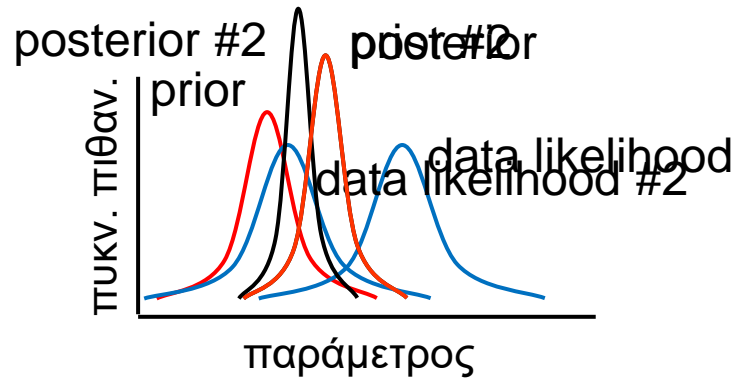
Εκτιμούμε τις παραμέτρους για κάθε dataset, οι οποίες είναι διαφορετικές κάθε φορά και υπολογίζουμε:

- Τα SD, που έχουν τη σημασία standard errors (SE),
- 95% CI κλπ
- Είναι εναλλακτικός τρόπος υπολογισμού των SE
- Δείχνει το robustness του μοντέλου

Empirical Bayes Estimates (EBE)

Θεώρημα του Bayes

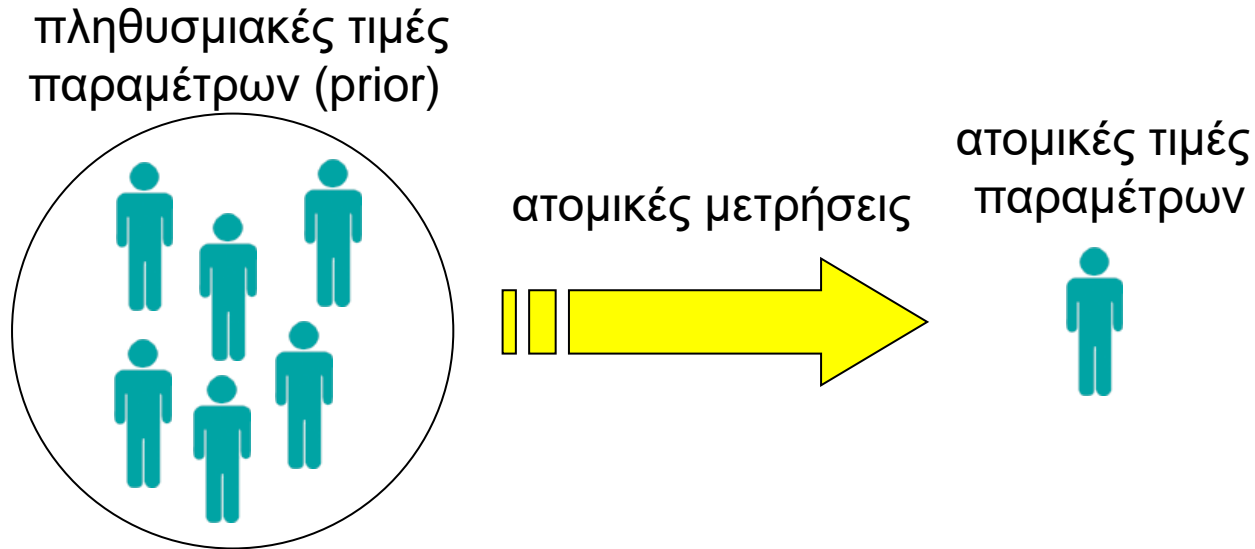
Το γινόμενο της **prior κατανομής** και της **πιθανοφάνειας των δεδομένων κανονικοποιημένο**, δίνει την **posterior κατανομή**



$$P(\theta|\mathbf{X}) = \frac{p(\theta) L(\mathbf{X}|\theta)}{\int p(\theta) L(\mathbf{X}|\theta) d\theta}$$

Διαδοχική εφαρμογή του θεωρήματος Bayes

Μπαεσιανή εξατομίκευση



- Εκτίμηση των ατομικών παραμέτρων των ασθενών
- Απαραίτητο σε κάποιες προσεγγίσεις του Likelihood function
- Χρήσιμο για covariate screening
- Χρήσιμο σε διαγνωστικά γραφήματα
- Εξατομίκευση της δόσης στην κλινική

Prior:

Γινόμενο κανονικών κατανομών.

Ένας όρος για κάθε παράμετρο

$$P(\xi) = \prod_{k=1}^p \frac{1}{\sqrt{2\pi\Omega_k}} e^{-\frac{(\xi_k - \theta_k)^2}{2\Omega_k}}$$

Likelihood:

Γινόμενο κανονικών κατανομών.

Ένας όρος για κάθε πειραματικό σημείο

$$L(C | \xi, \sigma) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(C_j - f(\xi, t_j))^2}{2\sigma_j^2}}$$

Θεώρημα Bayes

posterior \propto prior \times likelihood

$$L_{\text{posterior}} = \left\{ \prod_{k=1}^p \frac{1}{\sqrt{2\pi\Omega_k}} e^{-\frac{(\xi_k - \theta_k)^2}{2\Omega_k}} \right\} \times \left\{ \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(C_j - f(\xi, t_j))^2}{2\sigma_j^2}} \right\}$$

prior **likelihood**

Παίρνουμε το $-2\log(L_{\text{posterior}})$

$$F(\xi) = \sum_{k=1}^p (\theta_k - \xi_k)^2 / \Omega_k + \sum_{j=1}^n (C_j - f(\xi, t_j))^2 / \sigma^2$$

prior **likelihood**

Πώς γίνεται η Μπαεσιανή εξατομίκευση

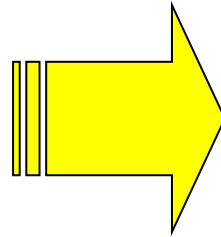
Οι εξατομικευμένες τιμές των παραμέτρων ξ , γνωρίζοντας τις πληθυσμιακές τιμές θ , Ω και σ , εκτιμώνται ελαχιστοποιώντας την αντικειμενική συνάρτηση:

$$F(\xi) = \underbrace{\sum_{k=1}^p (\theta_k - \xi_k)^2 / \Omega_k}_{\text{απόσταση των παραμέτρων από τις πληθυσμιακές τιμές}} + \underbrace{\sum_{j=1}^n (C_j - f(\xi, t_j))^2 / \sigma^2}_{\text{απόσταση του μοντέλου από τα δεδομένα}}$$

Όταν η παράμετρος ξ απομακρίνεται από την πλυθ. τιμή θ το F μεγαλώνει, τιμωρείται (penalty function)

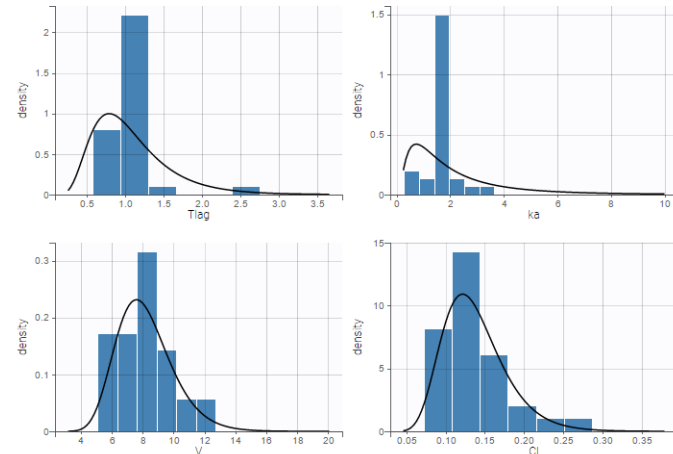
Population parameter estimates

Fixed Effects	
Tlag_pop	0.975
ka_pop	1.59
V_pop	7.98
Cl_pop	0.132
Standard Deviation of the Random Effects	
omega_Tlag	0.454
omega_ka	0.88
omega_V	0.221
omega_Cl	0.288
Error Model Parameters	
a	0.24
b	0.0603



Individual parameter estimates

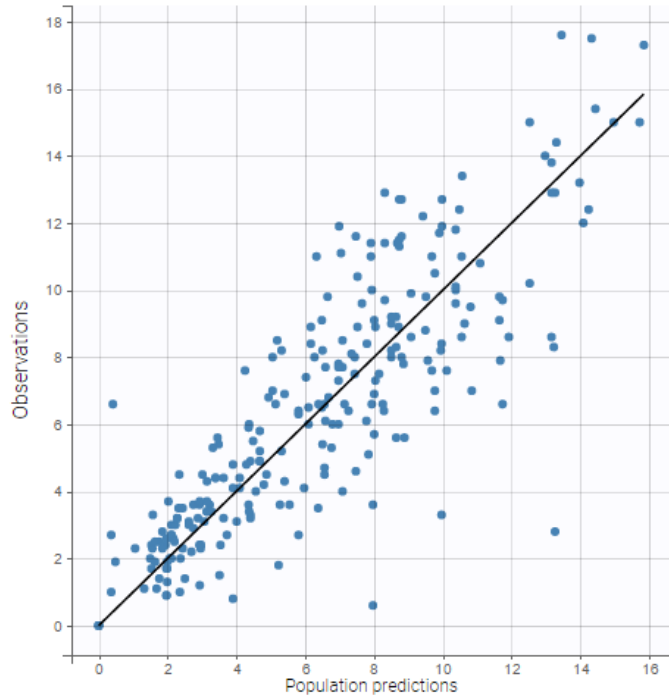
id	Tlag	ka	V	Cl
1	1.1200	2.7981	7.7367	0.1119
2	0.8408	1.1246	8.4987	0.1290
3	1.0536	2.4730	8.6165	0.1116
4	1.0187	3.1038	5.1427	0.0721
5	0.9987	2.0382	11.5424	0.1768
6	0.9618	3.3574	6.2530	0.2287
7	0.8046	5.2274	8.1075	0.1719
8	0.9609	0.6065	5.3130	0.1091
9	1.0734	2.7144	9.1817	0.1156



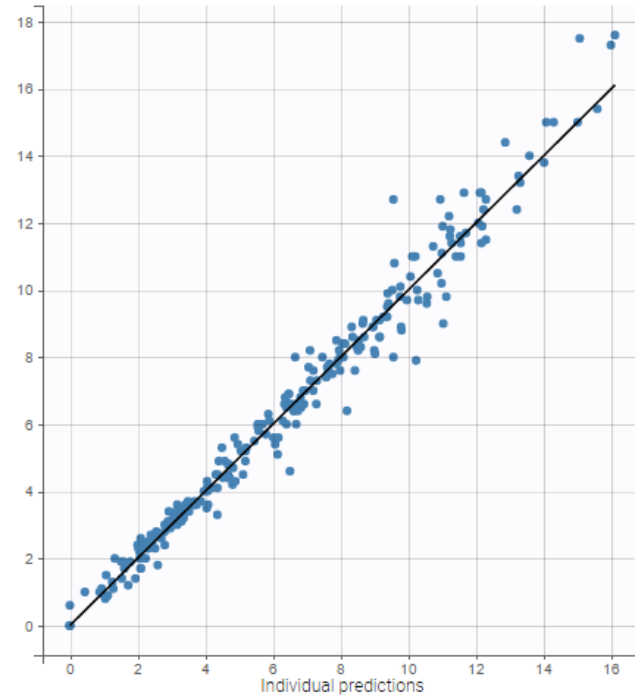
Διαγνωστικά γραφήματα με ΕΒΕ

OBS vs PRED

Population



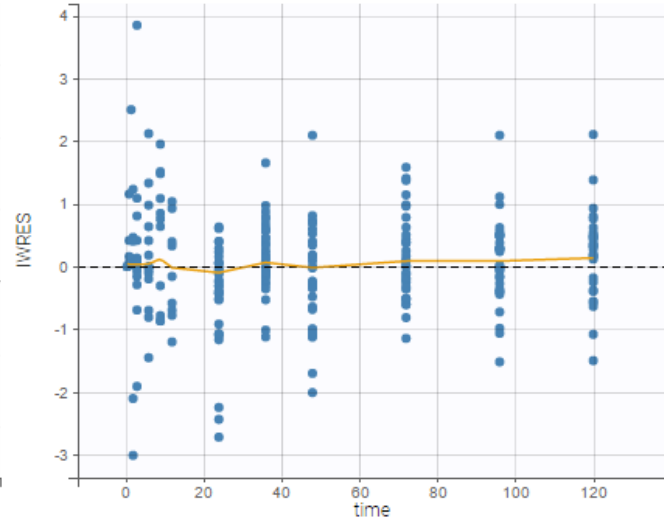
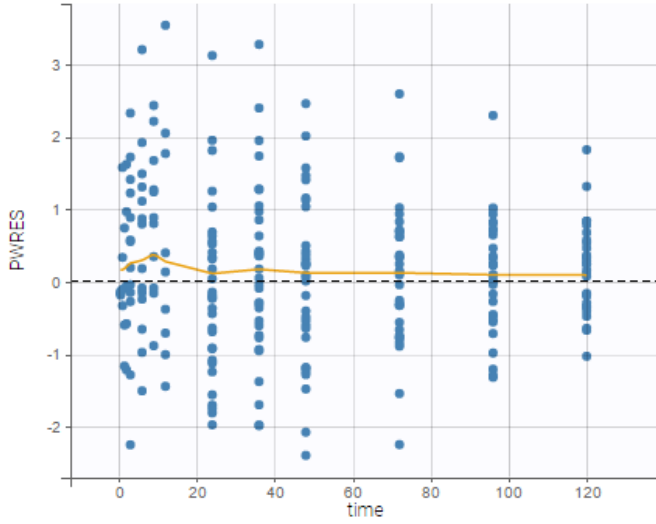
Individual



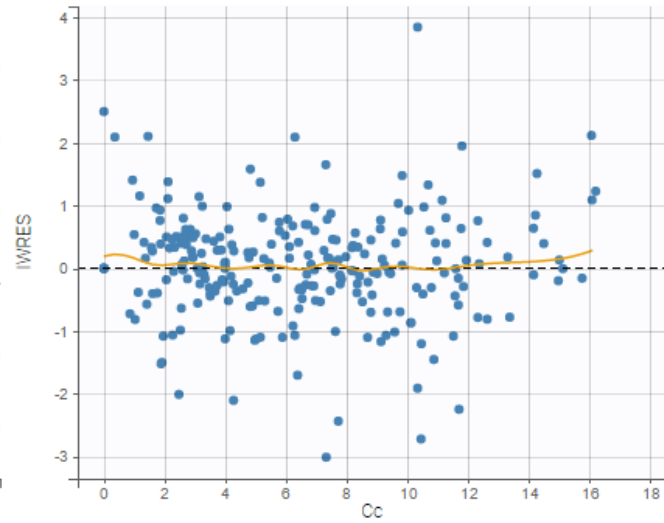
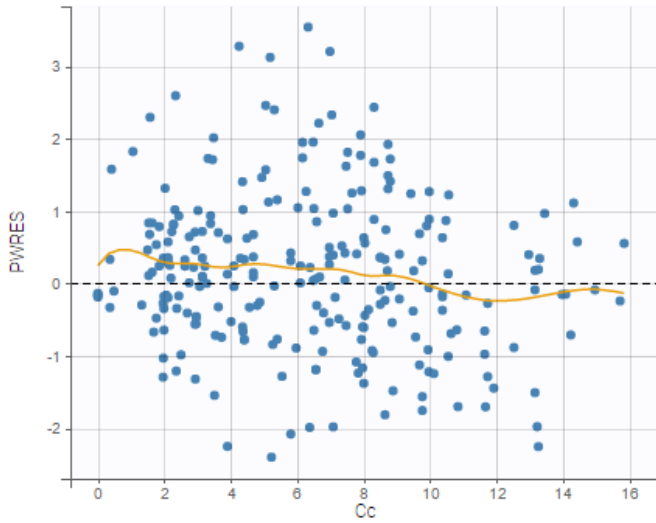
Population

Individual

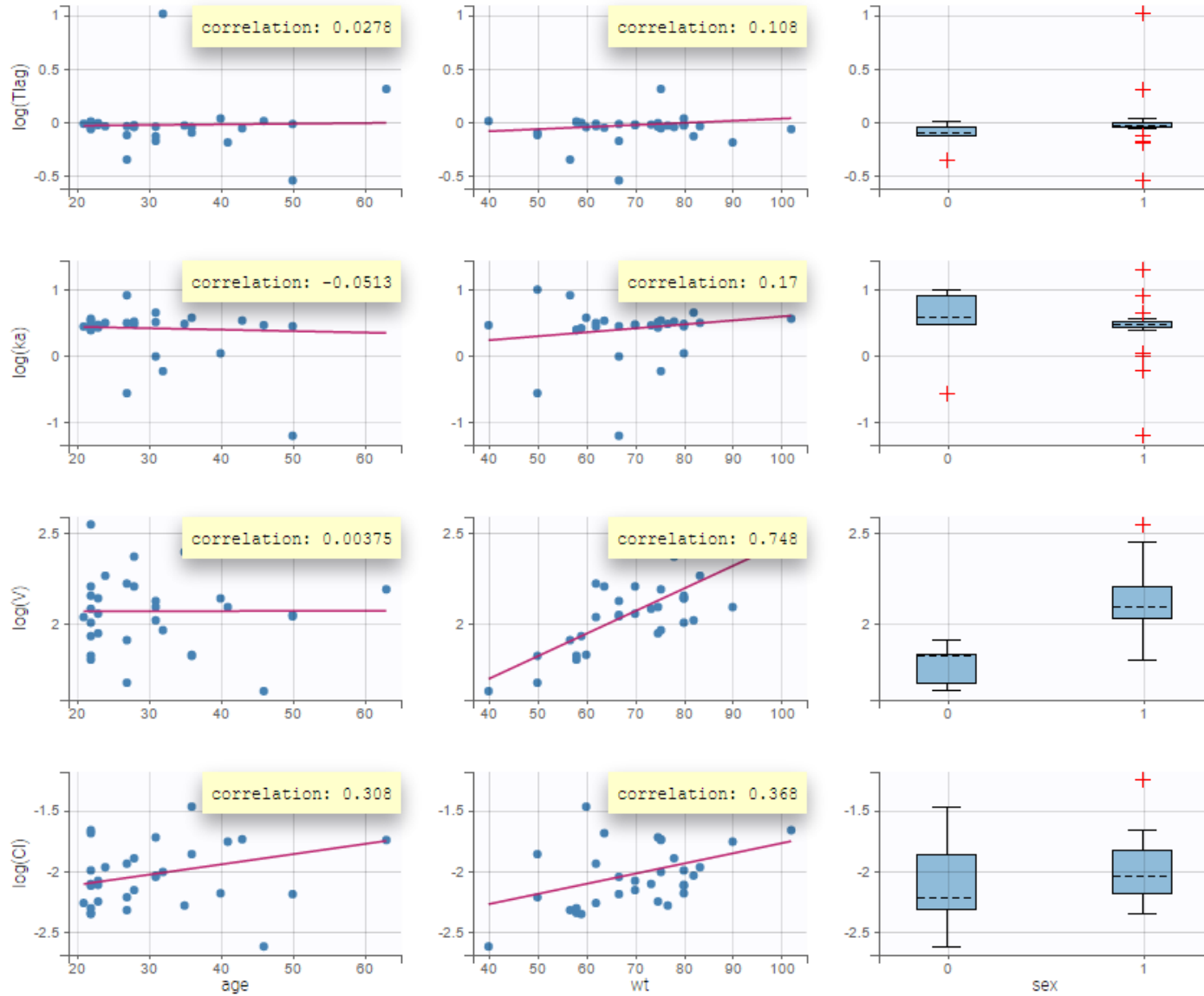
RES vs time



RES vs PRED

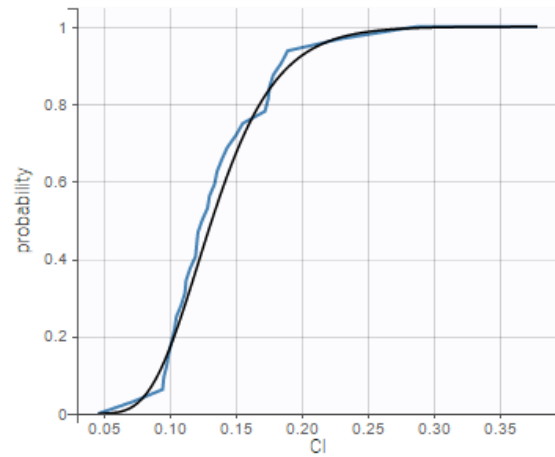
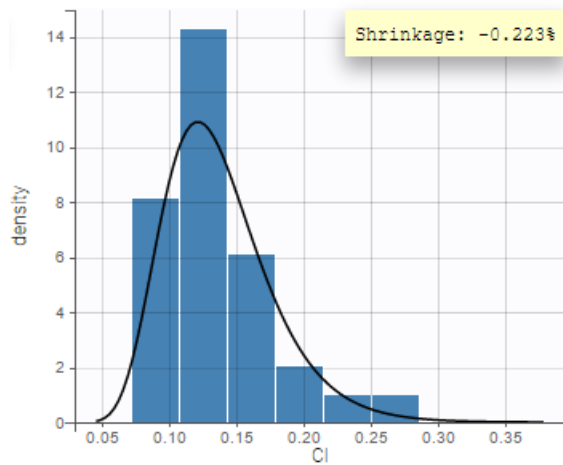
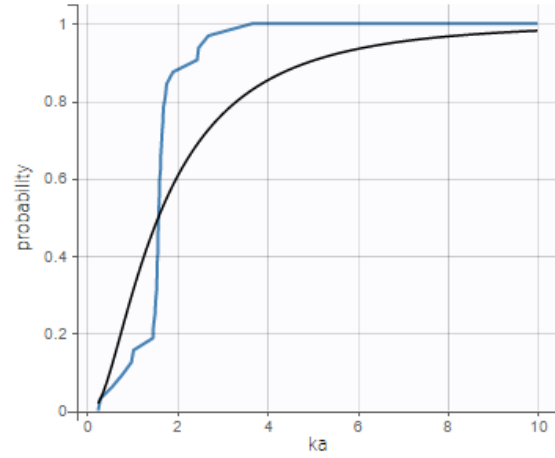
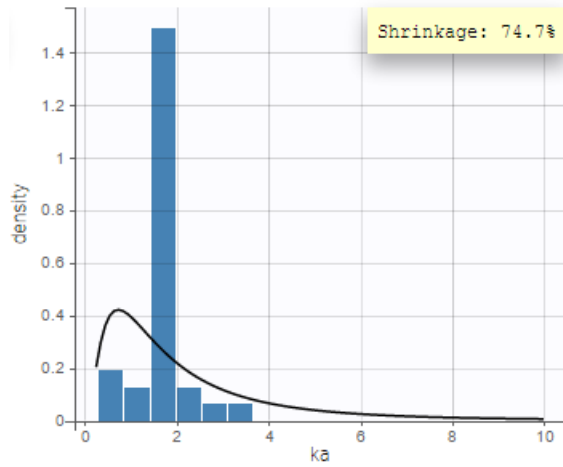


covariate screening $\mu\epsilon$ EBE

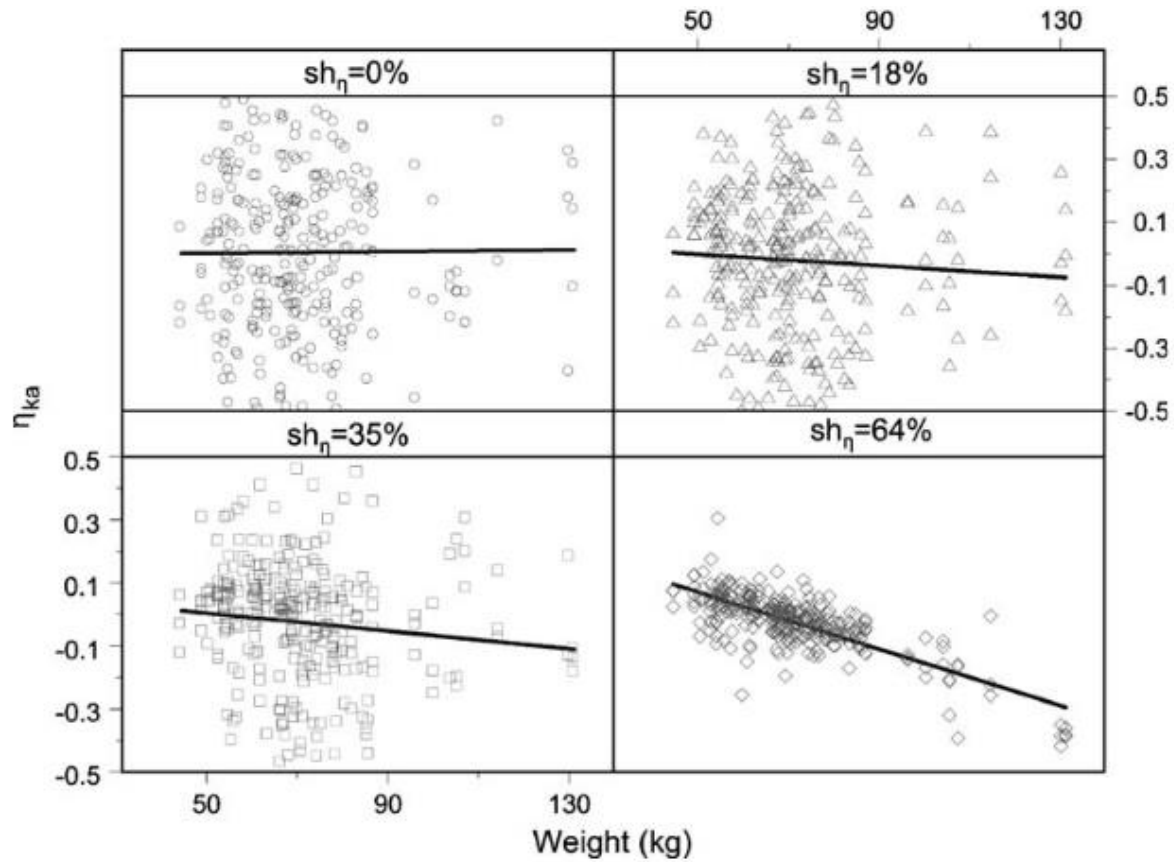


Shrinkage

Eta-shrinkage $sh_{\eta} = 1 - \frac{SD(\eta_{EBE})}{\omega}$

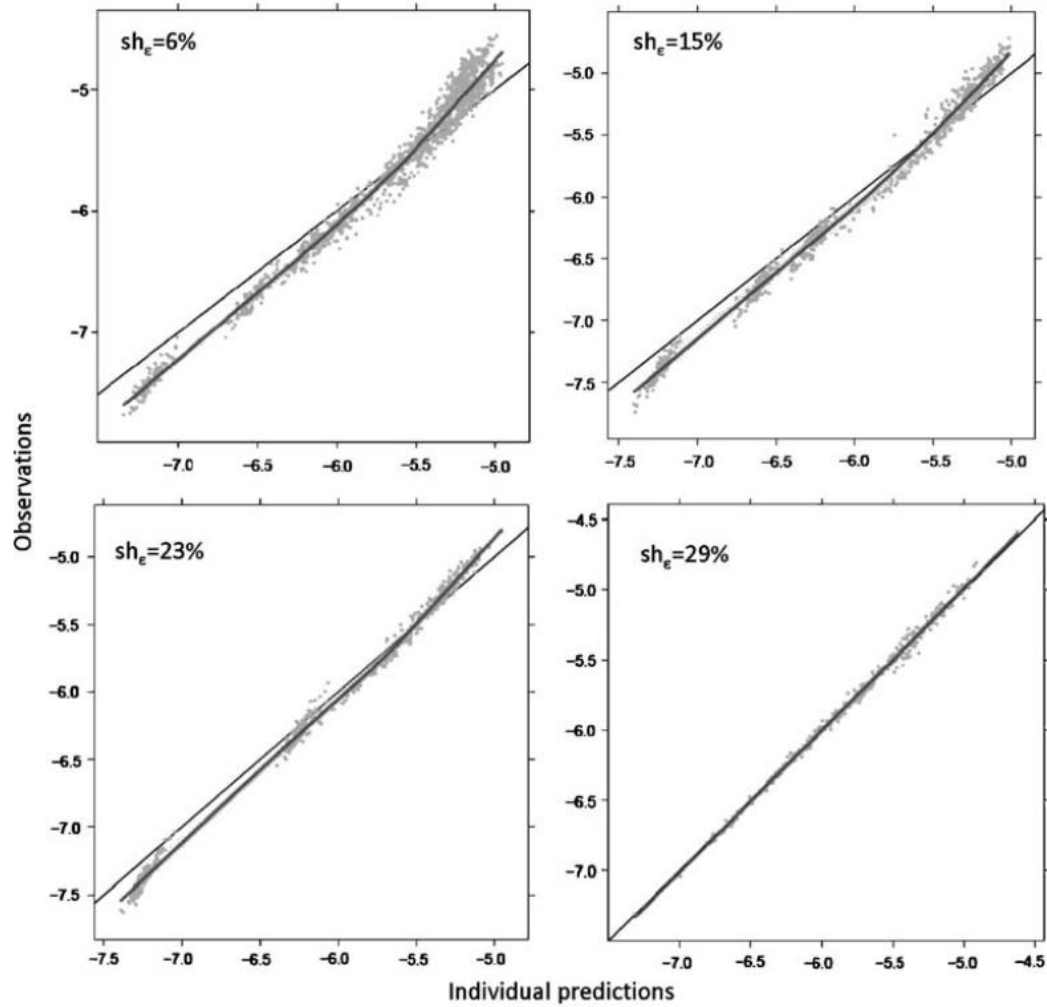


Επίδραση του η -shrinkage στο covariate screening



Epsilon-shrinkage
Perfect fit phenomenon

$$sh_{\epsilon} = 1 - SD(IWRES)$$



Επισκόπηση της διαδικασίας ανάπτυξης πληθυσμιακού μοντέλου

- **Κατασκευή σετ δεδομένων**
- **Ανάπτυξη Base model**
 - Exploratory analysis
 - δομικό μοντέλο ✓
 - μοντέλο σφάλματος ✓
 - Διατομική μεταβλητότητα IIV (αρχικά χωρίς συσχετίσεις) ✓ ✓
 - Inter-occasion variability IOV (αν χρειάζεται) ✓
 - Έλεγχος συσχετίσεων στην IIV
 - Εισαγωγή συσχετίσεων στην IIV ✓
 - Επικύρωση Base model ✓ ✓
- **Ανάπτυξη covariate model**
 - Έλεγχος συσχετίσεων ETA με covariate
 - Εισαγωγή covariate στο μοντέλο ✓
 - Αξιολόγηση covariate (στατιστική, φυσική, κλινική)
 - Επανάληψη για άλλον covariate
 - Πιθανή επανεξέταση του IIV ✓
- **Επικύρωση τελικού μοντέλου** ✓ ✓

- | | |
|---|--|
| ✓ | Έλεγχος GOF, SE, shrinkage, φυσική σημασία estimates |
| ✓ | ✓ Όλα τα παραπάνω + Επικύρωση με VPC, bootstrap |

Software - αλγόριθμοι

Η likelihood function στα Nonlinear Mixed effect models δεν εκφράζεται αναλυτικά

Τα διάφορα λογισμικά υλοποιούν είτε αριθμητικές προσεγγίσεις, είτε γραμμικοποίηση που επιτρέπει αναλυτική έκφραση

NONMEM, (Icon plc, αρχικά UCSF από το 1980)

Δύσχρηστο πρόγραμμα σε FORTRAN με τη λογική του configuration file

Αρχικά FO method (γραμμικοποίηση) , αργότερα FOCE, τώρα πλέον σχεδόν όλες τις διαθέσιμες μεθόδους, όπως Expectation Maximization, MCMC.

Monolix, (Lixoft) μοντέρνο user interface

Stochastic Approximation Expectation Maximization method (SAEM)

Phoenix, Certara, μοντέρνο user interface

FOCE και Expectation Maximization

Stan, (Stan Development Team) configuration file,

αντικατέστησε το BUGS/WinBUGS

Bayesian MCMC method

Software: Monolix, (Lixoft)

Algorithm: SAEM

Stochastic Approximation Expectation Maximization

Download link:

<https://lixoft.com/download/win64-monolix-suite-2023r1/>

Προς το τέλος της εγκατάστασης που ζητήσει license
πατήστε το κουμπί OK για 10 day free trial

