Analytical Performance Specifications: the benefit for the laboratory!

Haverkate lecture
10th ECAT Participants´meeting
Sverre Sandberg, Noklus / EFLM
Bergen, Norway

thebmj

BMJ 2015;351:h6431 doi: 10.1136/bmj.h6431 (Published 3 December 2015)

NEWS

Of the
values
were l
discor
impor

Data on trial of anticoagulant is to be reanalyzed after discovery that investigators used faulty device

Deborah Cohen

The BMJ

Data from the key phase III trial that allowed the oral anticoagulant rivaroxaban on to the market for the prevention of ischemic stroke in non-valvular atrial fibrillation is being reanalyzed because of the use of a faulty device.

The European Medicines Agency has also told *The BMJ* that it has launched an investigation into the trial.

The ROCKET-AF trial, published in the *New England Journal of Medicine* in 2011, compared rivaroxaban, which is verified by laboratory tests. Alere told the Project on Government Oversight that it was working on a software upgrade to fix the inaccuracies.

Even before the recall, the ROCKET-AF trial had been criticized for the relatively short time that participants taking warfarin were in the correct therapeutic range of the drug. The implication of the criticism was that patients' use of warfarin was not adequately controlled, which, if correct, would have the effect

BACKGROUND

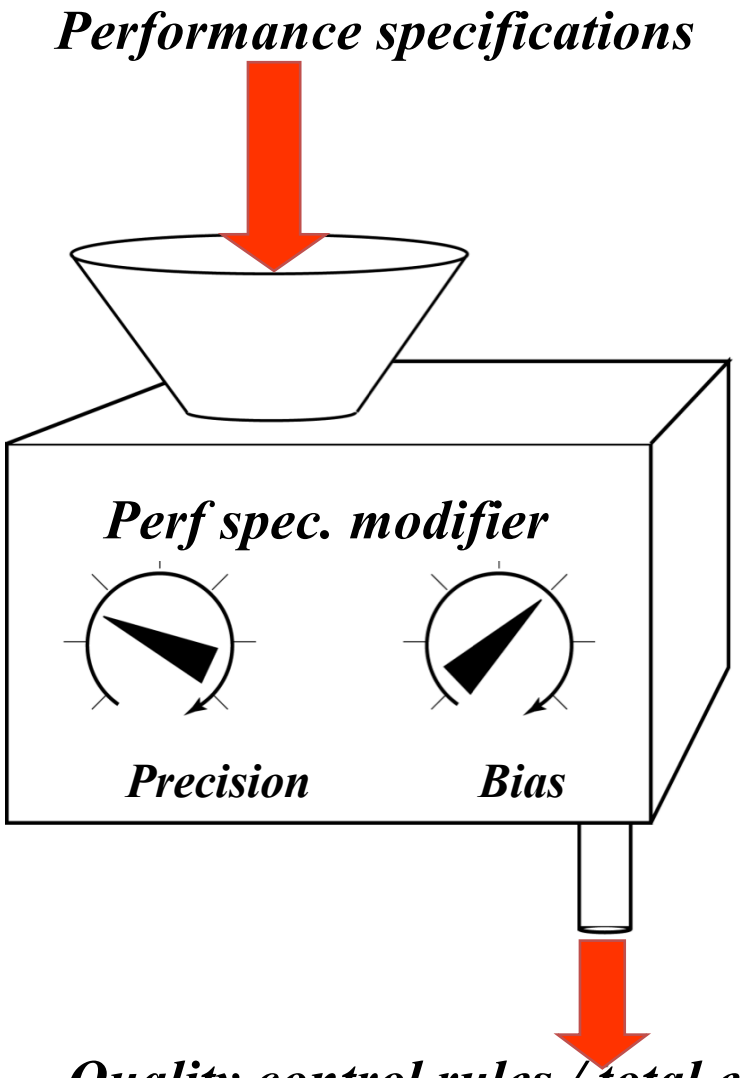
The use of warfarin reduces the rate of ischemic stroke in patients with atrial fibril-

From the Duke Clinical Research Institute

NOKLUS

this case. They note that diagnostic devices "rarely require prospective clinical studies for clearance" and that devices that have been recalled can be used as predicates for the approval of new devices. Even the FDA acknowledges that "there are few performance standards… based on clearly defined scientific parameters."

The importance of Performance specifications



EFLM
EUROPEAN FEDERATION OF CLINICAL CHEMISTRY AND LABORATORY MEDICINE

European Commission
Joint Research Centre
IRMM
Institute for Reference Materials and Measurements

CIRME
Università degli Studi di Milano

1st EFLM Strategic Conference

Defining analytical performance goals 15 years after the Stockholm Conference

8th CIRME International Scientific Meeting

Milan (IT)
24-25 November 2014

with the auspices of IFCC International Federation of Clinical Chemistry and Laboratory Medicine

Consensus statement

DE GRUYTER

Clin Chem Lab Med 2015; 53(6): 833–835

Consensus Statement

Sverre Sandberg*, Callum G. Fraser, Andrea Pi... ...sen, Graham Jones, Wytze Oosterhuis, Per Hyltoft Petersen, Hein... ...is and Mauro Panteghini

Defining analyti... ...iance specifications: Consensu... ...t from the 1st Strategic Confere... ...he European Federation of Clinical Chemist... ...y and Laboratory Medicine

CCLM 2015;53 issue 6

Remember that

George Box 1919-2013

"All models are wrong, but some are useful."

"The best models are not necessarily the most useful models".

NOKLUS

Model 1. Based on the effect of analytical performance on clinical outcomes

1a. Direct outcome studies

1b. Indirect outcome studies

Model 2. Based on components of biological variation of the measurand

Model 3. Based on state of the art

NOKLUS

Model 1. Based on the effect of analytical performance on clinical outcomes

This can, in principle, be done using different types of studies:

Direct outcome studies – investigating the impact of analytical performance of the test on clinical outcomes;

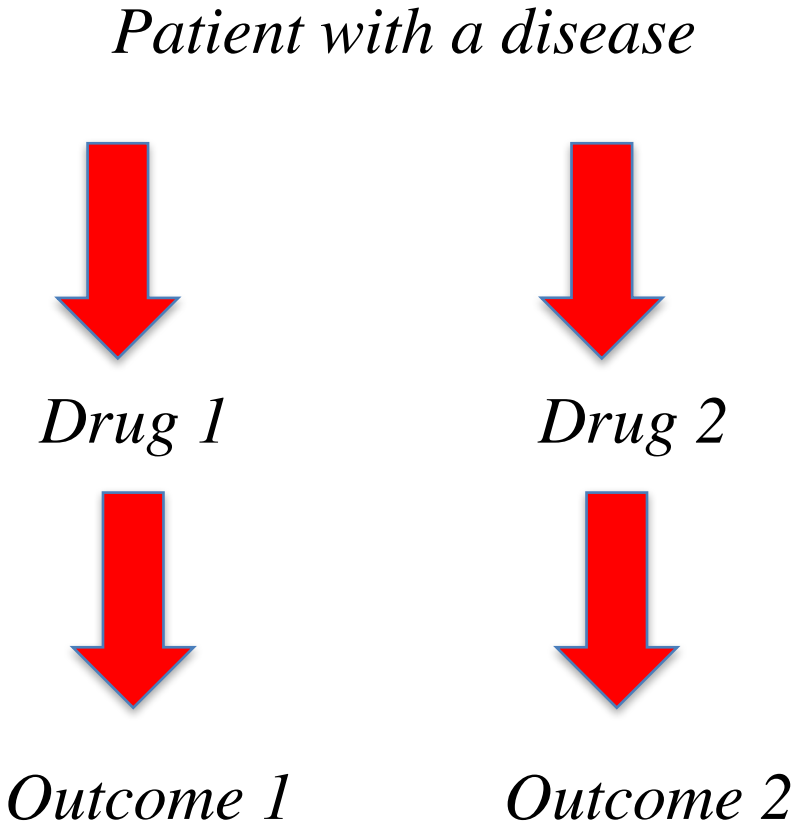
Indirect outcome studies – investigating the impact of analytical performance of the test on clinical classifications or decisions and thereby on the probability of patient outcomes, e.g., by simulation or decision analysis.

NOKLUS

The advantage of this approach is that it addresses the influence of analytical performance on clinical outcomes that are relevant to patients and society. The primary disadvantage is that it is only useful for examinations where the links between the test, clinical decision making and clinical outcomes are straightforward and strong.

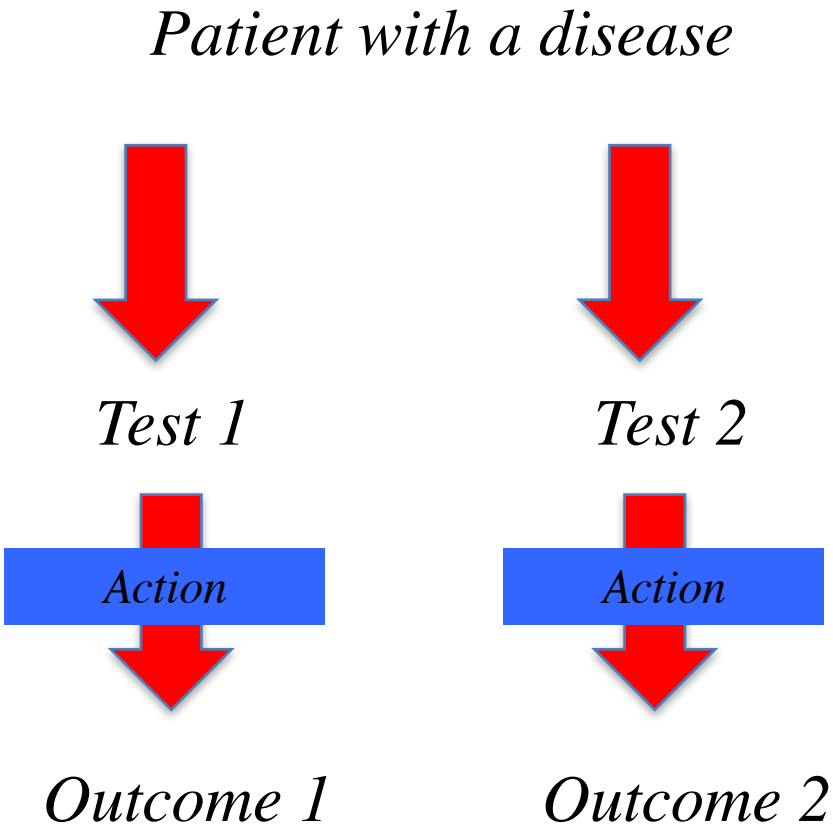
NOKLUS

Clinical outcome



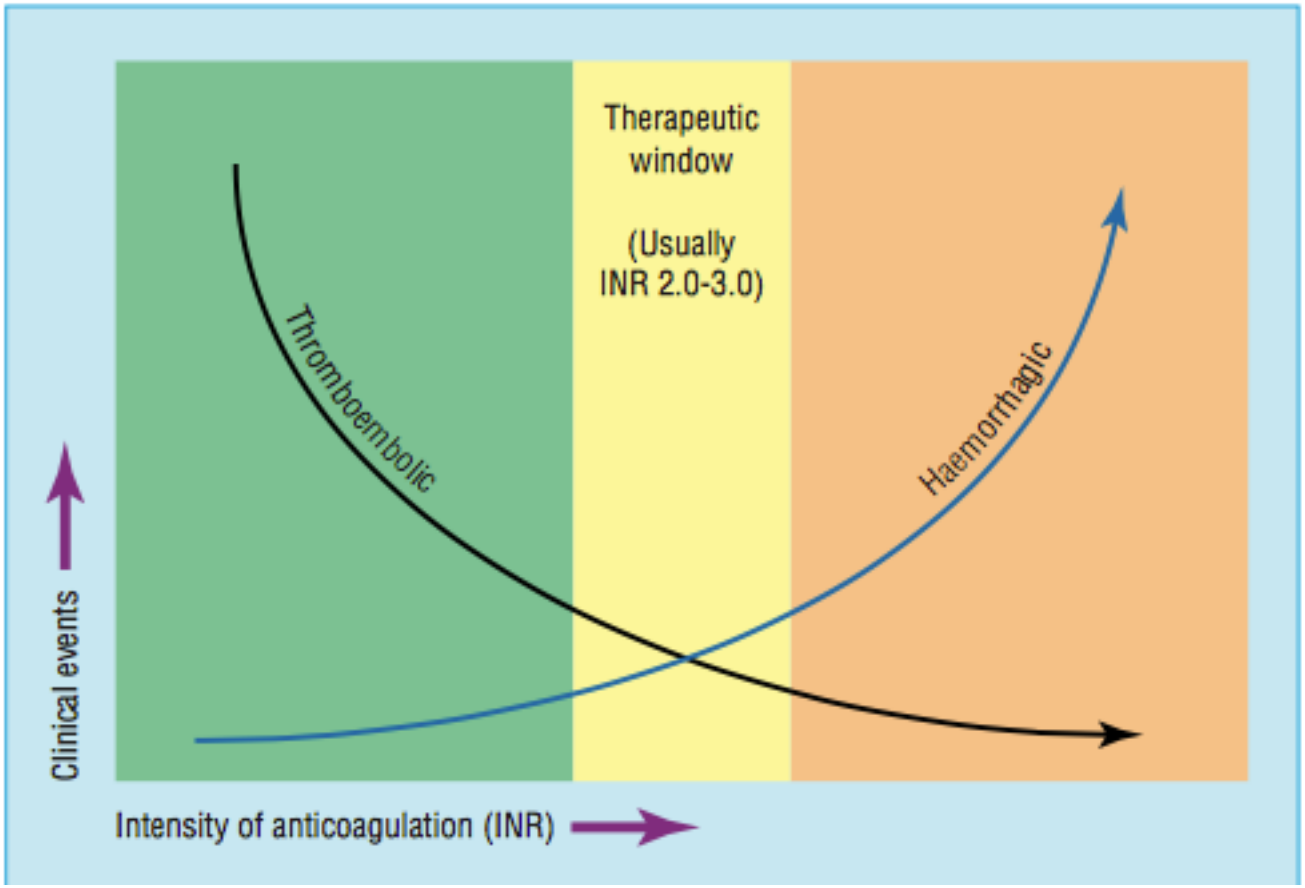
NOKLUS

Clinical outcome



NOKLUS

1b:
INR – balance between complications



So Model, type 1 is very difficult to perform and will take a lot of resourses.

However type 1b is not that difficult – but we still have to concentrate on one clinical condition – for example to see how many false positive of false negative we can accept

NOKLUS

Model 2. Based on components of biological variation of the measurand

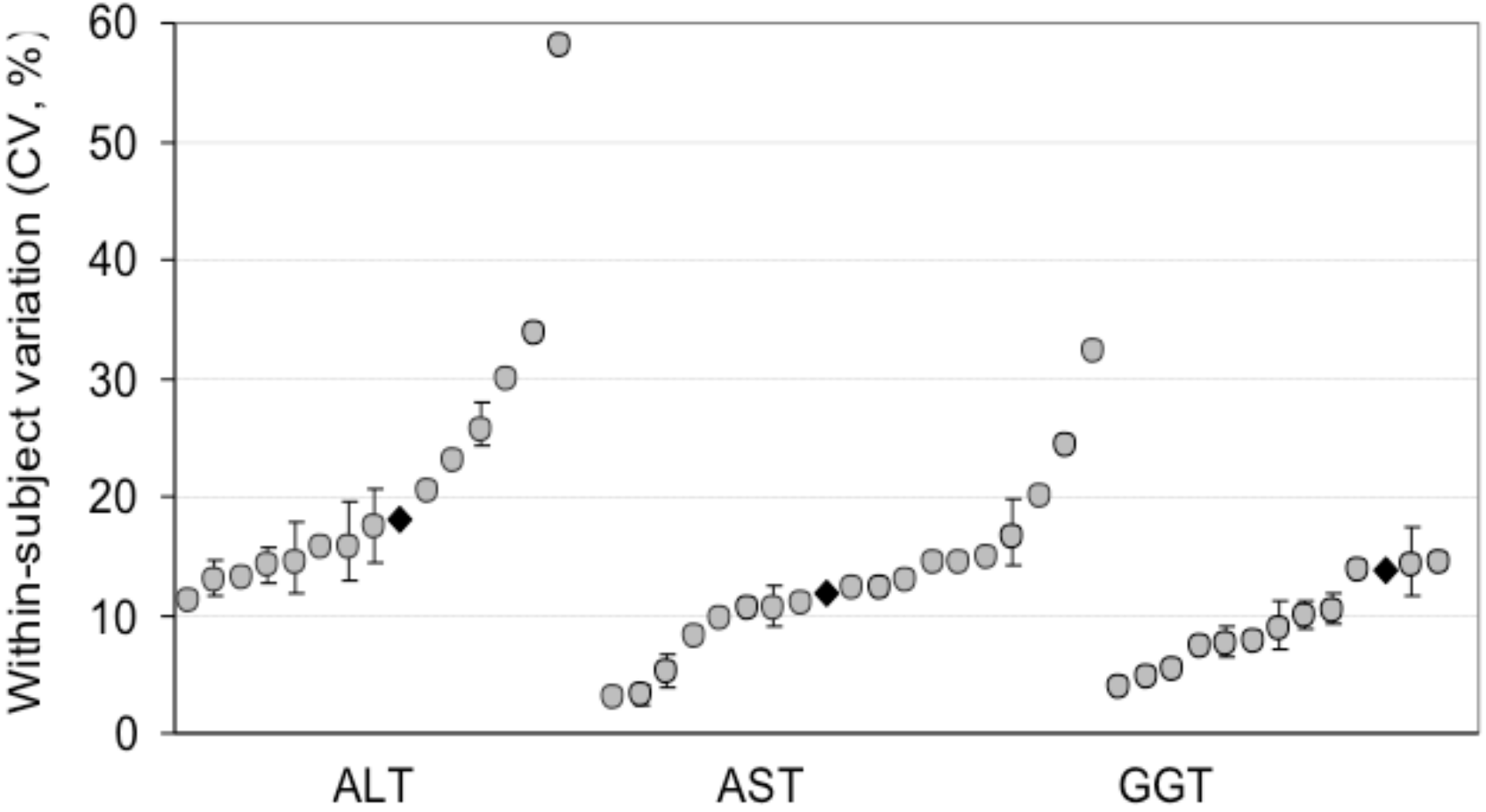
This attempts to minimize the ratio of 'analytical noise' to the biological signal.

The *advantage* is that it can be applied to most measurands for which population based or subject-specific biological variation data can be established.

The *limitations* are that much of the current data/studies have not been carried out in a proper scientific way and therefore contains flaws.

NOKLUS

Within-subject variation



Carobene A et al Clin Chem Lab Med. 2013;51:1997–2007.

NOKLUS

Model 3. Based on state-of-the-art

This relates to the highest level of analytical perfor- mance technically achievable. Alternatively, it could be defined as the analytical performance achieved by a certain percentage of laboratories.

The advantage of this model is that state-of-the-art performance data are readily available. The disadvantage is that there may be no relationship between what is technically achievable and what is needed to minimize the ratio of ‘analytical noise’ to the biological signal or needed to obtain an improved clinical outcome.

NOKLUS

Explanatory notes

- It should be noted that the three models use differ- ent principles.
- The hierarchy assumes that high quality studies or data are available for each model.
- Proposed analytical performance specifications should therefore always be accompanied by a statement of the rationale, the source and the quality of the evidence behind the recommendation.

NOKLUS

Same measurand – different performance specifications?

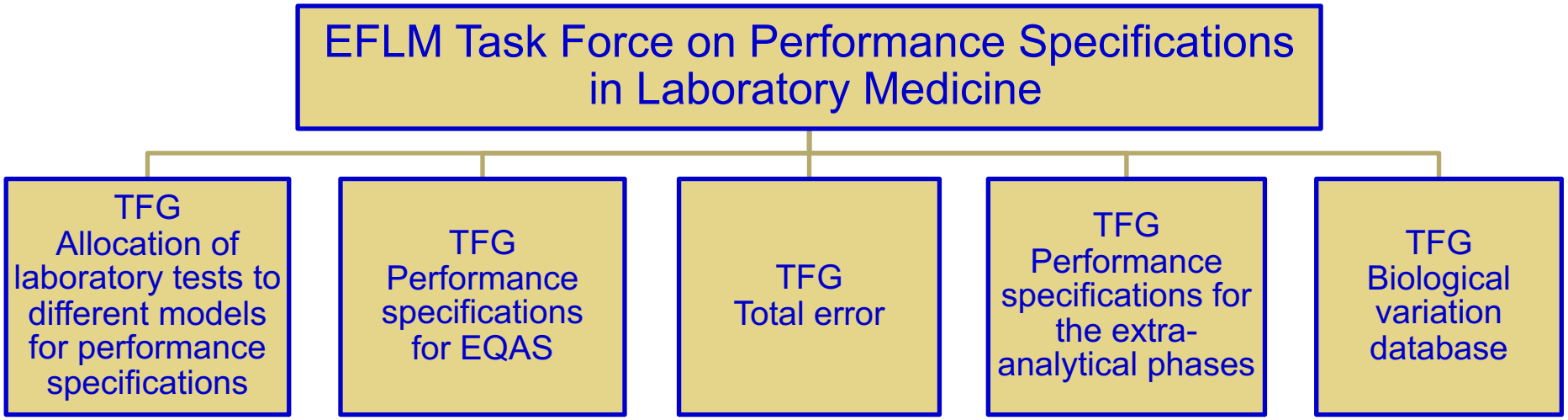
Some measurands can have different performance specifications defined when the test has multiple intended clinical applications.

This is difficult – if not impossible – for the laboratory to handle.

NOKLUS

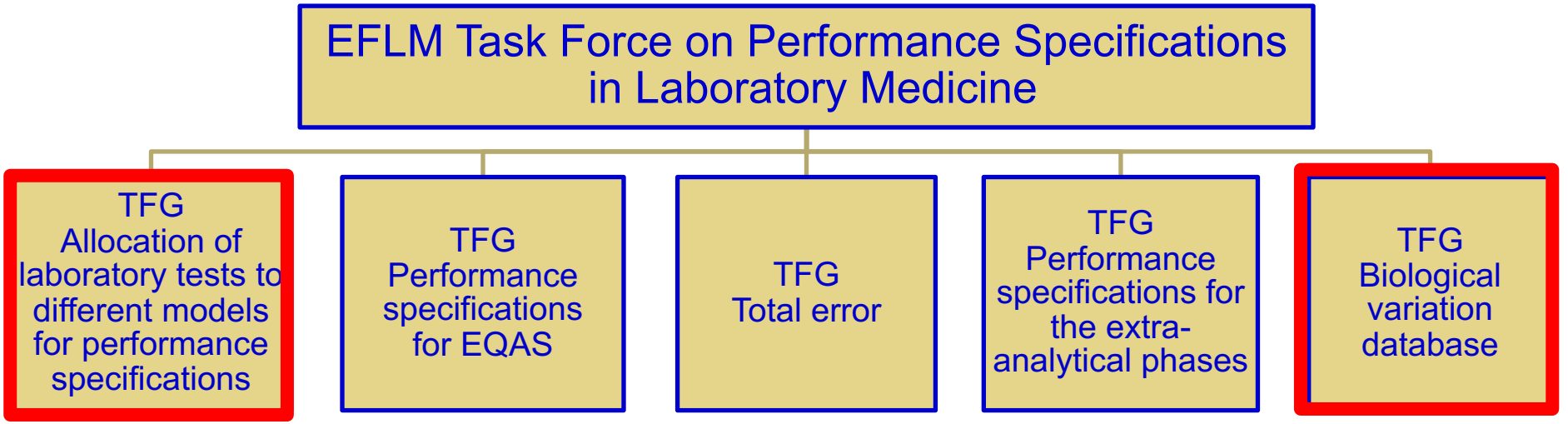
EFLM Task Force on Performance Specifications in Laboratory Medicine

An **EFLM Task Force on Performance Specifications in Laboratory Medicine (TF-PS)** has been created to coordinate the activities of the Task & Finish Groups (TFG) established as outcome of the 1st Strategic Conference



EFLM Task Force on Performance Specifications in Laboratory Medicine

An **EFLM Task Force on Performance Specifications in Laboratory Medicine (TF-PS)**
has been created to coordinate the activities of the Task & Finish Groups (TFG) established as outcome of the
1st Strategic Conference



TFG: Allocate tests to different models

Possible principles:

Model 1 - outcome

- Measurands that are mainly used for one purpose e.g. HbA1c / INR – can be used as general PS - in EQAS for example
- Specific purposes e.g. Selfmonitoring INR, glucose, kreatinin. Can be used in guidelines for specific situations but not as general PS

Ferruccio et al. Clin Cehm Lab Med 2016, in press

NOKLUS

Model II – biological variation

- Measurands in which it is important for the body to keep a constant concentration- for example electrolytes, Hb, some hormones

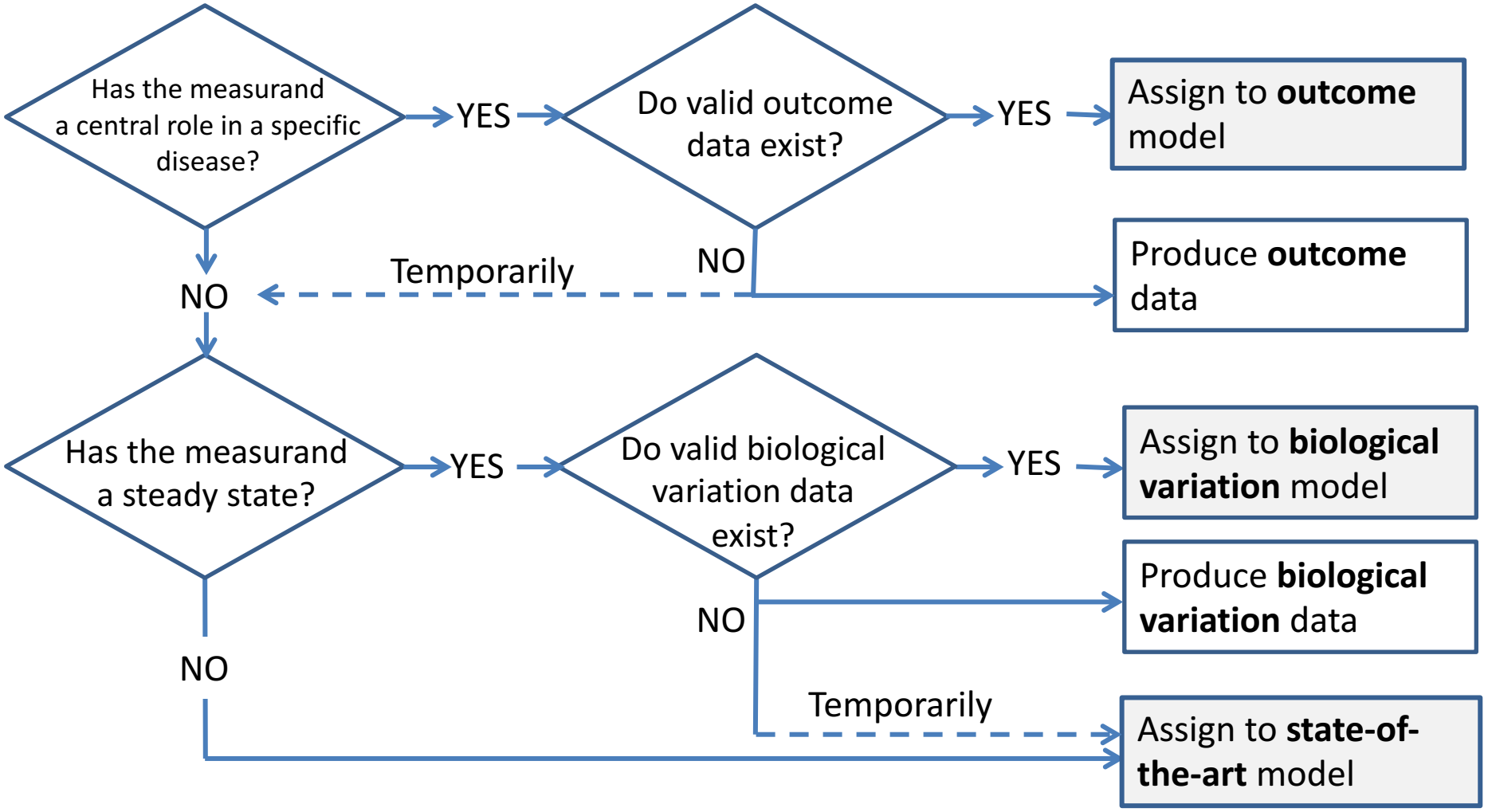
Model III – state of the art

- Measurands where it is possible to obtain data from model I or model II;

Ferruccio et al. Clin Cehm Lab Med 2016, in press

NOKLUS

Model assignment workflow



Ferruccio et al. Clin Cehm Lab Med 2016, in press

Biological variation database
- cooperation -

Callum Fraser / Per Hyltoft Petersen

Analaytical Quality comission in SEQC

EFLM – TFG

35 persons working with papers for kidney disease, cardiac markers, lipids, kidney diseases, diabetes.

NOKLUS

An improved biological variation database

1. To use a critical appraisal check list to evaluate papers on biological variation.
2. To collect samples to analyze most measurands on biological variation
3. To collect data on biological variaton in a new database on www.eflm.eu

NOKLUS

First report from the EFLM Task and Finish Group (TFG) for the Biological Variation Database (BVD)

By Sverre Sandberg
Chair of the EFLM TFG on Biological Variation

Critical appraisal check list

- ✓ Papers are categoirzed as A, B, C and D depending on their methodological quality, with category A papers indicating high-quality and D poor quality.
- ✓ The checklist contains 14 items, and
- ✓ 22 items will be extracted from each paper and presented in the database.
- ✓ Groups are established for different measurands

NOKLUS

“EFLM BV-WG project: new experimental BV data Samples collection from healthy volunteers for biological variation values update”

A multicenter study involving 6 European laboratories

[3] Haukeland University Hospital, Bergen, Norway.

[2] Wilhelmina Ziekenhuis Assen Europaweg-Zuid 1, 9401 RK Assen, the Netherlands.

[4] Hospital Universitario La Paz, Madrid, Spain, and Quality Analytical Commission of Spanish Society of Clinical Chemistry (SEQC).

[5] Dept. of Laboratory Medicine University Hospital, Padova –Italy.

[1] Servizio Medicina di Laboratorio, Ospedale San Raffaele, Milano, Italy.

[6] Acibadem University, School of Medicine, Atasehir, Istanbul, Turkey.

Carobene et al.
Clin Chem Lab Med, in press

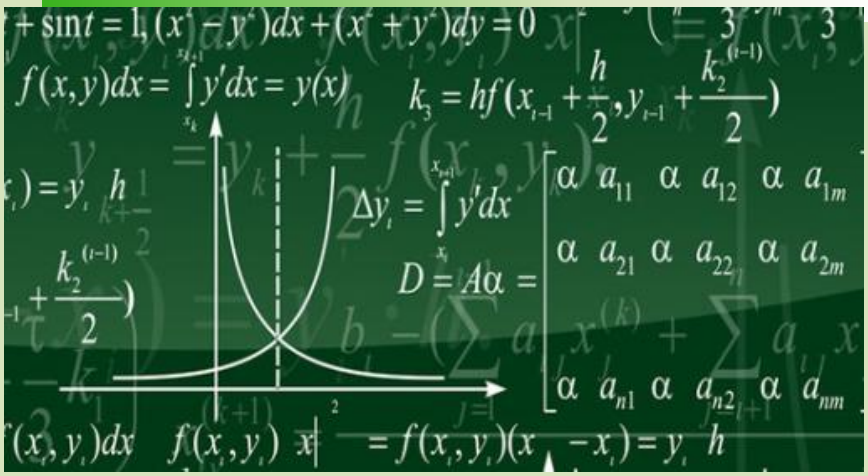
Methods: Each lab has processed and frozen the specimens strictly in agreement with the procedure reported in the protocol, in order to avoid pre-analytical variability.

Results: A total of about 19.500 aliquots have been collected: 120 aliquots of serum, 40 of plasma EDTA, and 40 of plasma citrate for each subject. The samples, delivered to the coordinator lab (Milan, Italy), are stored at -80 ºC in a dedicated freezer.

Carobene et al. Clin Chem Lab Med, in press

The data will be treated according to recommended procedures for calculating BV including: outliers analysis at three different levels (replicate, samples and subjects), homoscedasticity of variances, normality of the distribution, and a nested ANOVA.

Future work:
A large number of tests, including enzymes, substrates, proteins, electrolytes, hormones, vitamins, tumor markers and coagulation tests will be performed in duplicate in the same analytical run.



When we have good studies: How to use the data?

When people say they use ”biological variation” ”clinical opinion” to set performance specifications, there is a big variation.

Why?

Because they use different (sub)models

Different levels of models

- Principal models – what principles should we use to set performance specifications
- (The Milan confererence)

- Practical models – when we have chosen the principles – which models should we then use to set practical performance specifications
- (A lot of work to be done)

Second level models

First choose what of the three principal models to use,

then choose a ”submodel”

Performance specifications - what are you interested in?

Bias

Imprecision

Total error

Measurement uncertainty

Outcome based models

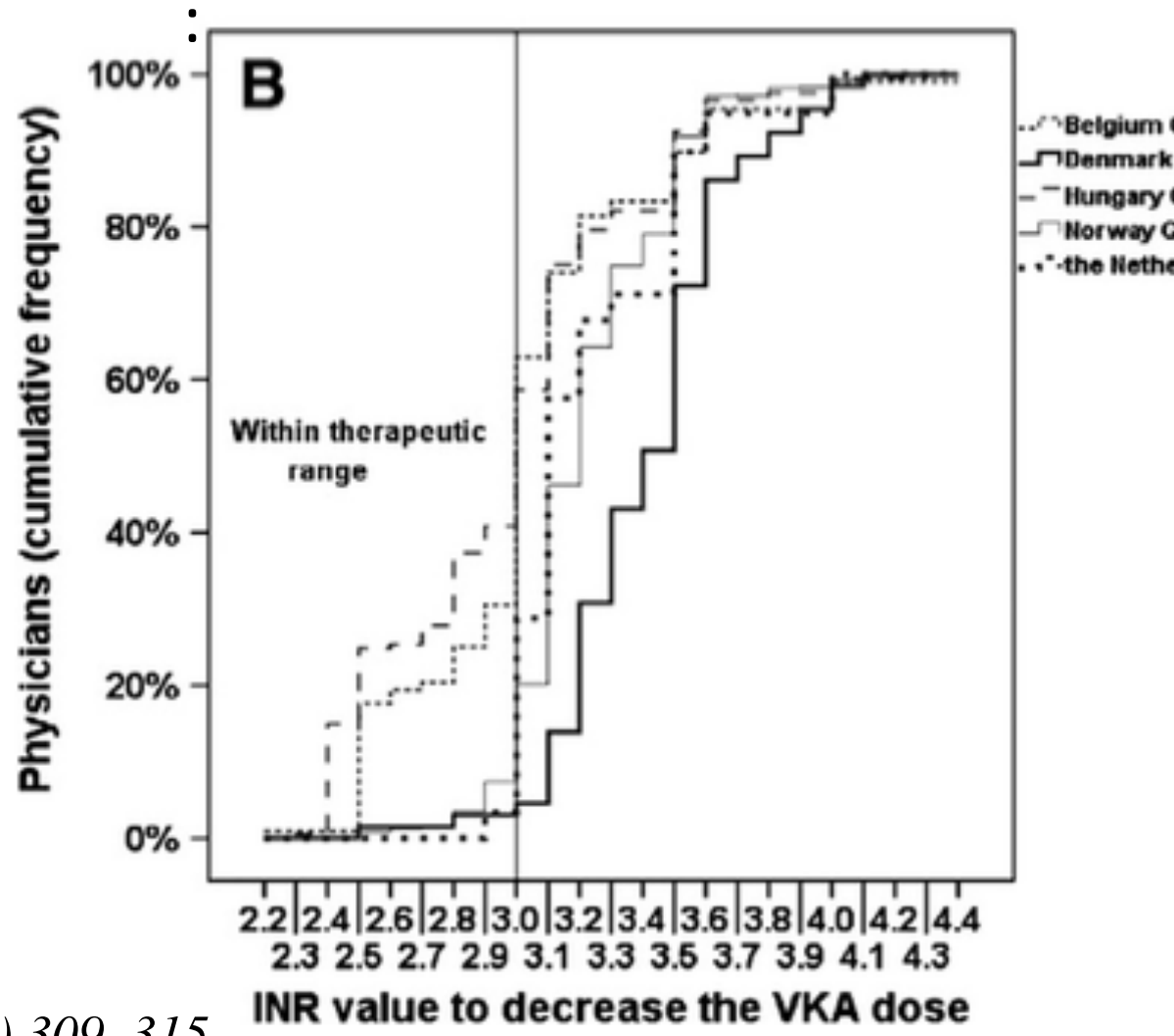
If outcome data should be used – should then the most strict data be used or the less strict?

- For clinical guidelines it will be easier to use outcome based specifications since they usually deal with one specific clinical indication.

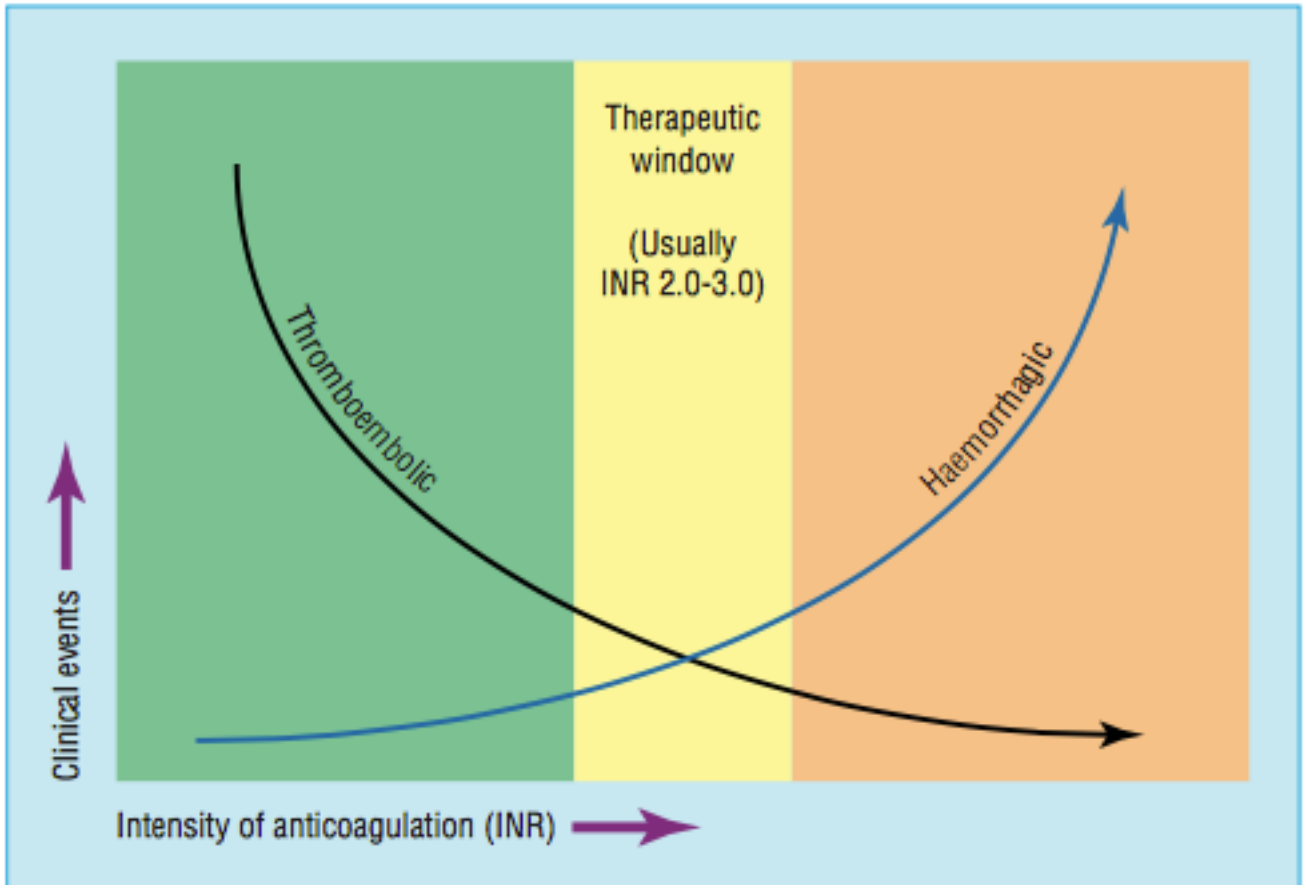
Example INR: Clinical algorithms for dosing - based on experience

- Indirect experience with analytical quality of measurement
- Frequency of dosing
- Amount of drugs given
- Fixed limits (e.g. target limits)

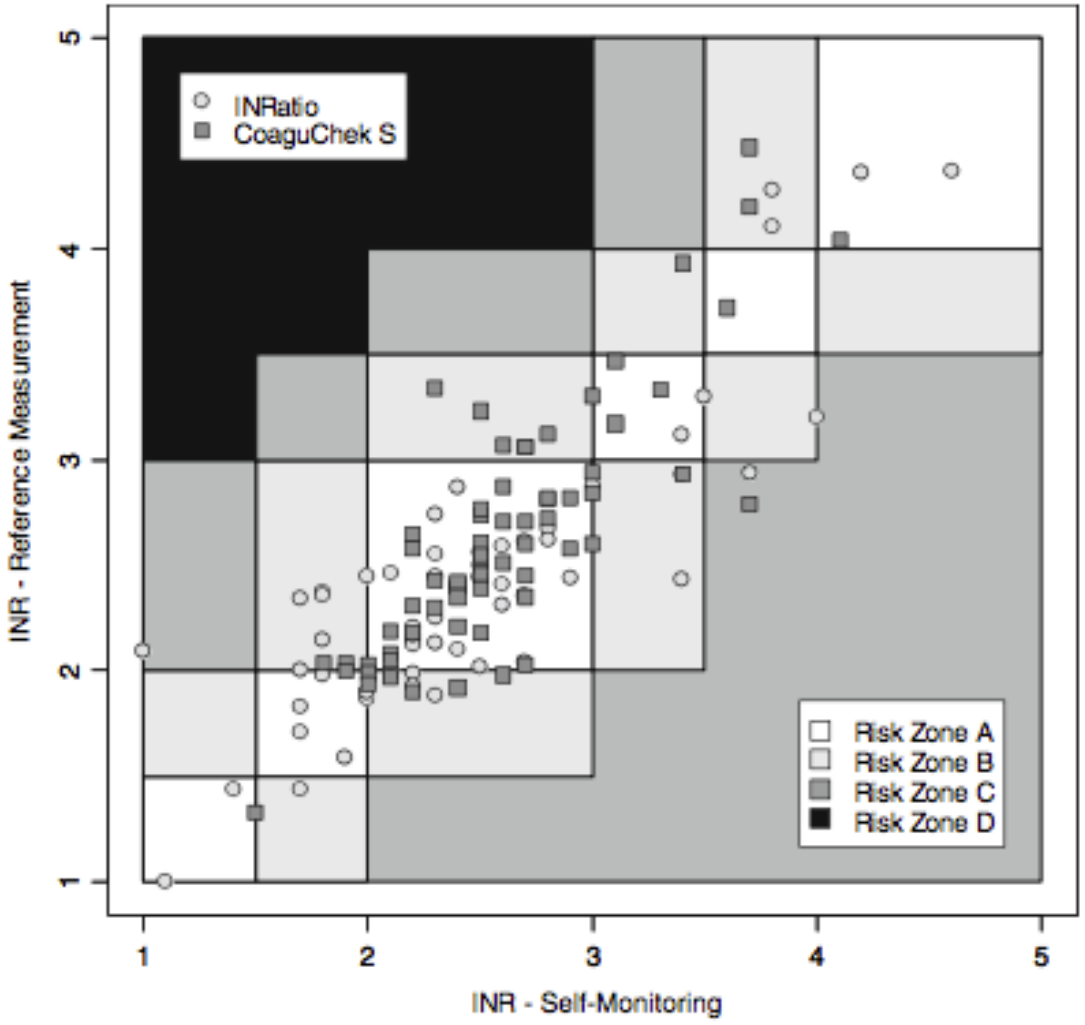
Patient with atrial fibrillation and an INR of 2.3.
INR values (representing critical differences) at which physicians would increase or decrease the VKA dose.



Thrombosis Research 130 (2012) 309–315


Therapeutic window
(Usually INR 2.0-3.0)
Thromboembolic
Haemorrhagic
Clinical events
Intensity of anticoagulation (INR)

Model 1b: Error grid



J Thromb Thrombolysis (2008) 26:22–30

Submodels for biological variation

- PS for monitoring
- PS for diagnosing
- PS for EQAs
- PS for method evaluation

Models based on biological variation - relating to the reference interval -

Tonks: “TE in %” ≤ ± 100∗(0.25 times the “normal range”

Harris: $s_A^2 + B^2 \leq 0.25 * (s_I^2 + s_G^2)$

Gowans: to keep the common ref intervals within the uncertainty of a 120 person sample. Instead of the usual 2.5% outside a reference limit a maximum of 4.6% (based on the IFCC guide- line on reference values)

$|B_{MAX}| \leq 0.25 * s_B = 0.25 * (s_I^2 + s_G^2)^{½}$,

SD_A max < 0.52 * $(s_I^2 + s_G^2)^{½}$,

Pragmatic model where you add analytical CV to within-subject or between subeject variation

(Oosterhuis et al/ Hoetzel et al – clin chem lab med 2015)

Table 4. CV_A as measured in this study in relation to the recommen

Variable	CV_A: this study	Desirable CV_A: this study		Desirable CV_A: lit	
		Diagnosis	Monitoring	Diagnosis	
General assays					
PT	1.3	2.8	1.3	3.0–4.2	1.2–1.3 [C
APTT	1.4	5.5	3.3	5.0–6.4	0.9–2.3 [C *(14)*, Ch et al. *(21*
Fibrinogen	2.7	12.0	6.0	10.0–13.5	2.6–6.7 [d *(13)*, Ch et al. *(16*
Thrombosis factors					
AT	1.4	5.2	2.2	4.4–6.1	0.6–2.7 [C et al. *(17*
PC-clot	5.5	14.0	4.5	7.8–11.1	1.2–4.0 [C *(18)*]
PC-chrom	1.7	12.1	3.8	–	
PS-act	4.1	14.6	4.0	13.7	3.8 [Wada
		11.1	3.7	5.4	1.5 [Nguy

De Maat et al Clin Chem 2016 in press

$CV_A \leq 2\%$	Ricos et al. 1999 (healthy individuals)
$CV_A \leq 1.5\%$	Wada et al 2004 (healthy individuals)
$CV_A \leq 1.65\%$	Rudez et al 2009 (healthy individuals)
$CV_A \leq 4.5\%$	Lassen et al. 1995 (patients on OAT)
$CV_A \leq 4.5\%$	Besselaar et al. 2012 (patients on OAT)
$CV_A \leq 4.6\%$	Geest-Daalderop et al. 2009 (patients on OAT)
$CV_A \leq 5.4\%$	Geest-Daalderop et al. 2010 (patients on OAT)
$CV_A \leq 1,3\%$	De Maat et al 2016 (healthy individuals)

INR Biological variation

$CV_A = 0,5\ CV_I$

There is an important relationship between

analytical performance (bias and imprecision)

and

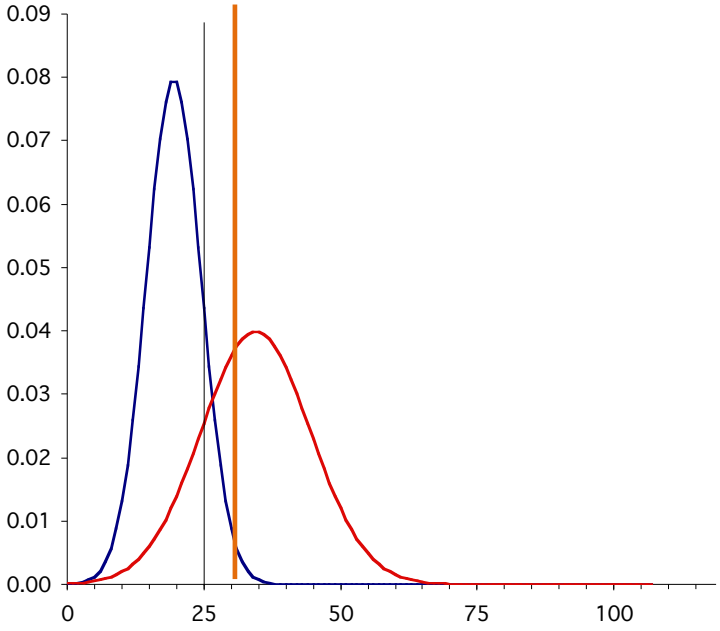
clinical performance (sensitivity and specificity / predictive values)

Effect of analytical bias

Bias	0.0

	ND	D
Mean	20.0	35.0
SD	5.0	10.0

Cut-off	30
Sens	0.69
Spec	0.98



Conclusions

- Three main models using different principles
- Many ”submodels” to calculate performance specifications
- Biological variation most used in laboratories and EQAS
- There is a profund relationship between analytical performance and clinical performance

Thank you