Til deltakerne på "The Artic Experience 2016":

SKUPs rapporter er ganske omfattende.
Dette er et utdrag av rapporten,
og dekker den informasjonen som er nødvendig
for å delta «Finse-kurset» 2016

microINR portable coagulometer

A system for measurement P—Prothrombin time (INR)
manufactured by iLine Microsystems S.L.

Report from the evaluation SKUP/2015/109

organised by SKUP at the request of iLine Microsystems S.L.

SKUP in Denmark, Dept. KBA, Nordsjællands Hospital, 3400 Hillerød, Phone +45 4829 4176, www.skup.nu
SKUP in Norway, Noklus, Box 6165, NO-5892 Bergen, Phone +47 55979502, www.skup.nu
SKUP in Sweden, Equalis, Box 977, SE-751 09 Uppsala, Phone: +46 18 490 31 44, www.skup.nu

2. Abbreviations and Acronyms

BLS	Biomedical Laboratory Scientist
CI	Confidence Interval
C-NPU	The committee on Nomenclature, Properties and Units
CV	Coefficient of Variation
DEKS	Danish Institute of External Quality Assurance for Laboratories in Health Care
EQA	External Quality Assessment
Equalis	External quality assurance in laboratory medicine in Sweden
NKK	Norwegian Clinical Chemistry EQA Program
Noklus	Norwegian Quality Improvement of Primary Care Laboratories
NS_EN ISO/ IEC	Norsk Standard_Europeisk Norm International Organization for Standardization/International Electrotechnical Commission
PHCC	Primary health care centre
PT (INR)	Prothrombin Time International Normalized Ratio
RBT	Rabbit Brain Thromboplastin
SKUP	Scandinavian evaluation of laboratory equipment for primary health care
WHO	World Health Organization

3. Introduction

3.1. Background for the evaluation

The microINR portable coagulometer is produced by iLine Microsystems S.L. and was launched into the Scandinavian market December 2012. iLine Microsystems S.L. is the requesting company in this evaluation. A first SKUP evaluation of the microINR system was organized during spring 2014. This evaluation was stopped halfway on the request from the producer, due to underperformance of an R&D Software version provided for the evaluation.

3.2. The aim of the evaluation

The aim of the evaluation was to determine the analytical quality and user-friendliness of the microINR system, both when used under optimal conditions in a hospital laboratory and by intended users in primary health care.

The evaluation includes:

- Examination of the analytical quality (precision and accuracy) under optimal conditions
- Examination of the analytical quality (precision and accuracy) in the hands of intended users
- Evaluation of the user-friendliness of the microINR system and manual

3.3. The SKUP model

SKUP evaluations for quantitative methods are based upon the fundamental guidelines in a book concerning evaluations of laboratory equipment in primary health care [1]. A complete SKUP evaluation consists of two parts. One part of the evaluation is carried out under optimal conditions in a hospital laboratory. This part documents the quality of the system under conditions as favourable as possible for achieving good analytical quality. The other part of the evaluation is carried out by intended users in at least two primary health care centres (PHCCs). This part documents the quality of the system under real-life conditions.

The evaluation under optimal conditions includes:

- Repeatability with 100 patient samples
- Comparison with an established hospital laboratory method

The evaluation performed by intended users includes:

- Repeatability with 40 patient samples at each of the primary health care centres
- Comparison with an established hospital laboratory method
- Evaluation of user-friendliness

If possible, SKUP evaluations are carried out using three lot numbers of test chips from separate and time-spread productions.

4. Quality goals

4.1. Analytical quality

At the present, there are no generally recognised analytical quality goals for the determination of prothrombin time International Normalized Ratio (PT (INR)), and no international (Gold) standard for evaluation of Point of Care test instruments for PT (INR) in primary health care.

The International Organization for Standardization (ISO) 17593 standard [2] gives requirements for monitoring systems for self-testing of oral anticoagulant therapy. In SKUP's opinion, the quality goals for accuracy in the ISO 17593 standard (±30% for 90% of the results in the therapeutic range 2 – 4,5 PT (INR)) is too tolerant. Furthermore, there is no performance criterion for imprecision in the standard.

Setting quality goals based on biological variation is an acknowledged method [3,4]. It is recommended that analytical imprecision (repeatability, coefficient of variation CV_a) should be less than, or equal to, half the intra-individual biological variation. Ricos *et al*. [5] state the biological variation for prothrombin time as CV_{bw} 4% (intra-individual biological variation) and CV_{bb} 6,8% (inter-individual biological variation). According to Kjeldsen *et al.* [6], the "in-treatment within-subject biological variation" of PT (INR) is 10,1%. For systems used for monitoring, the analytical performance should aim at low imprecision compared to the within-subject biological variation. In principle, quality goals based on biological variation do not take into account clinical requirements.

A committee appointed by the National Ministry of Health in Denmark has specified the requirements of analytical quality for PT (INR) [7,8] to bias ≤6% and imprecision ≤5% for instruments used in primary health care, and bias ≤3% and imprecision ≤3% for hospital instruments. There is no separate goal for the total error in the Danish specifications; however, estimated CV for the matrix effect is defined and an allowable deviation is given in the control system.

Based on the given data on biological variation for prothrombin time, and the fact that PT (INR) devices used in primary health care are designed for *monitoring* prothrombin time, SKUP recommends that these instruments should achieve a repeatability (CV)≤5,0%. SKUP has not taken out a separate goal for bias, but a figure of 5% was used to calculate a quality goal for allowable deviation according to the model below.

In method evaluations and comparisons, the imprecision of the comparison method has to be taken into account. SKUP allows an imprecision (CV) of the comparison method up to 3%. In addition, inter-laboratory variation should be taken into the calculation of the allowable deviation, which SKUP has estimated to 3%.

When comparing two prothrombin time methods, especially when the methods represent two different measuring principles, certain sample specific errors can be assumed. SKUP has chosen to include a variation of 5% in the error model for calculation of allowable deviation.

Allowable deviation = |±bias| + 1,65 x $\sqrt{CV^2_{test\ method} + CV^2_{comparisonmethod} + CV^2_{between\ lab} + CV^2_{matrix}}$

= (5 + 1,65 x $\sqrt{25+9+9+25}$) = (5 + 13,6) ≈ ±19%

4.2. User-friendliness

The evaluation of user-friendliness is carried out by asking the evaluating persons to fill in a questionnaire divided into four subareas, see section 6.5.

Technical errors

SKUP recommends that the percentage of "tests wasted" caused by technical errors should not exceed 2%.

4.3. Principles for the assessments

To qualify for an overall good assessment in a SKUP evaluation, the measuring system must show satisfactory analytical quality as well as satisfactory user-friendliness.

4.3.1. *Assessment of the analytical quality*

The analytical results are assessed according to the quality goals set for the evaluation.

Precision

The decision whether the achieved CV fulfills the quality goal or not is made on a 5% significance level. The distinction between the ratings, and the assessment of precision according to the quality goal, are shown in table 1.

Table 1. The rating of precision

Distinction between the ratings	**Assessment according to the quality goal**
The CV is lower than the quality goal (statistically significant)	The quality goal is fulfilled
The CV is lower than the quality goal (not statistically significant)	Most likely the quality goal is fulfilled
The CV is higher than the quality goal (not statistically significant)	Most likely the quality goal is not fulfilled
The CV is higher than the quality goal (statistically significant)	The quality goal is not fulfilled

Trueness

The confidence interval (CI) of the measured bias is used for deciding if a difference between the two methods is statistically significant (two-tailed test, 5% significance level). The term trueness is related to the results achieved under optimal conditions. Proven systematic deviation of the results achieved by intended users will be discussed in relation to the bias found under optimal conditions.

Accuracy
The accuracy is illustrated in a difference plot with limits for the allowable deviation according to the quality goal. The fraction of results within the limits is counted. The accuracy is assessed as either fulfilling the quality goal or not fulfilling the quality goal.

Bias with three lots of test chips
Separate lot calculations are not performed. The results achieved with the three lots are included in the assessment of accuracy in the difference plots. If distinct differences between the lots appear, this will be pointed out and discussed.

4.3.2. *Assessment of the user-friendliness*
The user-friendliness is assessed according to the answers and comments given in the questionnaire (see section 6.5.). For each question, the evaluator must choose between three given ratings. The responses from the evaluators are reviewed and summed up. To achieve the overall rating "satisfactory", the tested equipment must reach the total rating of "satisfactory" in all four subareas of characteristics described in section 6.5.

Technical errors
The evaluating person registers error codes and technical errors during the evaluation. The fraction of technical errors is calculated and taken into account in connection with the assessment of the user-friendliness.

4.4. SKUP´s quality goals in this evaluation
As agreed upon when working on the protocol, the results from the evaluation of the microINR system are assessed against the following quality goals:

Repeatability (CV) .....….≤5,0%

Allowable deviation in the individual result
from the comparison method result.....<±20%

Required percentage of individual results
within the allowable deviation*.....≥95%

Fraction of technical errors .....≤2%

User-friendliness, overall rating.....Satisfactory

*If more than 1% of the results deviate more than ±25%, this will be pointed out and discussed

5. Materials and methods

5.1. Definition of the measurand

The Committee on Nomenclature, Properties and Units (C-NPU) describes clinical laboratory tests in a database [9]. In the NPU-database the specifications for the measurand in this evaluation are as shown in table 2.

Table 2. NPU-specifications

NPU code	Name of test according to NPU	Unit
NPU01685	P—Coagulation, tissue factor-induced; relative time(actual/normal; INR; IRP 67/40; proc.)	–
NPU21717	P—Coagulation, tissue factor-induced; rel.time(actual/norm; INR; IRP 67/40; II+V+VII+X)	–

The analytical test according to NPU01685 refers to measurements performed with the Owren method. The test is mainly determined by the concentration of the Vitamin K dependent coagulation factors II, VII and X. The analytical test according to NPU21717 refers to measurements performed with the Quick method. The test is mainly determined by the concentration of the Vitamin K dependent coagulation factors, in addition to fibrinogen (factor I) and factor V.

Even if the tests according to NPU01685 and NPU21717 are not measuring exactly the same plasma components, the test results are used as if they did. In this report, the comparison method is an Owren method while the evaluated method, the microINR system, is a Quick method. The term "PT (INR)" will be used for the measurand in this report. As the measurement result is a ratio of the actual coagulation time divided with the normal coagulation time (INR), there is no unit.

5.2. The evaluated measurement system microINR

The information in this section derives from the company information material.

The microINR system is intended for the quantitative determination of PT (INR). The microINR meter (figure 1) and the disposable analytical microINR test chips compose the microINR portable coagulometer. The product is intended for professional use, self-testing and self-management of patients on oral vitamin K antagonist therapy.

Figure 1. microINR

The sample material on the microINR system is fresh capillary whole blood. The Rightest GD500 Lancing device is supplied for sampling. This device is for repeated use on a single person only and it was not used in this evaluation.

The microINR chips have two channels; one for sample measuring and one for control. Each micro capillary channel consists of a reaction chamber, where the reagent is placed, and a micro capillary, where PT (INR) is determined. The measuring channel contains dried reagent of human recombinant thromboplastin in a phospholipidic matrix and stabilizers, and so does the control channel, but in addition it contains human clotting factors.

Blood is inserted into the chip from the entry channel; it is divided into the two channels and mixed with the reagents, which activates the coagulation cascade. When blood clotting has occurred, PT (INR) calculation is performed from the monitored curves.

The calibration of the test is done automatically as the International Sensitivity Index (ISI) and Mean Normal Prothrombin Time (MNPT) value for the lot of the microINR chips is coded in the data matrix printed on each chip.

The manufacturer iLine Microsystems S.L. produces the control material microINR Easy Control. The material is lyophilised human abnormal plasma with buffer, stabilizers and preservatives.

According to the manual of the microINR system, there are two modes for sample application:
A: Approach finger to the meter
B: Approach the meter to the pricking area
This evaluation was performed using mode A, as instructed by the producer and supplier.

For technical details about the microINR system, see table 3. For more technical information about the microINR system, name of the manufacturer and the suppliers in the Scandinavian countries, see attachment 2 and 3. For product specifications in this evaluation, see attachment 4.

Table 3. Technical details from the manufacturer

Technical details for the microINR system	
Sample material	Fresh capillary blood
Sample volume	3 µL
Measuring time	
Measuring range	0,8 – 8,0 INR
Haematocrit	25% – 55%
Storage capacity	199 results
Electrical power supply	Rechargeable battery

5.3. The selected comparison method

A selected comparison method is a fully specified method, which in the absence of a Reference method, serves as a common basis for the comparison of the evaluated method.

5.3.1. *The selected comparison method in this evaluation*

The selected comparison method in this evaluation is the routine method for PT (INR) at the Department of Medical Biochemistry at St. Olavs Hospital in Trondheim, Norway, hereafter called "the comparison method".

The method is accredited after NS_EN ISO/IEC 17025 (Norsk Standard_Europeisk Norm International Organization for Standardization /International Electrotechnical Commission, 2005).

Instrument:	STA-R Evolution, STAGO, two identical instruments were in use	
Reagent:	STA-SPA+, Diagnostica STAGO	
Principle:	Owren's method, rabbit brain thromboplastin and adsorbed bovine plasma	
Traceability:	World Health Organization's (WHO's) manual tilt tube technique and the reference thromboplastin WHO IRP 67/40, through Rabbit Brain Thromboplastin (RBT/90) [9-11]	
Calibrators:	Three point's calibration with PT (INR)-calibrators from Equalis (External quality assurance in laboratory medicine in Sweden)	
Reference range	0,9 – 1,2 INR	
Therapeutic range	venous indication	2,0 – 3,0 INR
	arterial indication	2,5 – 3,5 INR

Internal analytical quality control

Internal analytical quality control samples, two levels (STA-Scandinorm PT (INR) and STA-Scandipath PT (INR), STAGO) were measured each evaluation day on the comparison method instruments. The reproducibility of the comparison method as achieved with the quality control material was calculated.

External analytical quality control

The hospital laboratory participates in Noklus/NKK (Norwegian Quality Improvement of Primary Care Laboratories/Norwegian Clinical Chemistry EQA Program) external quality assessment (EQA) scheme for PT (INR) with two levels four rounds per year. The materials are freshly frozen pooled citrate plasma from Norwegian donors. The assigned value for PT (INR) is based on consensus values from 67−69 participants using PT (INR)-calibrators from Equalis.

5.3.2. *Verification of the analytical quality of the comparison method*

Precision

The repeatability of the comparison method was calculated from duplicate measurements of venous citrate samples from patients in stable (≥4 weeks) vitamin K antagonist treatment. The requested repeatibility of the laboratory is 2,4% at a PT (INR) level of approximately 3,0.

Trueness

- PT (INR) calibrators from Equalis were analysed as samples on the comparison method at the start and in the end of the evaluation. The calibrator material is a pool of citrated anti-coagulated freeze-dried plasma of human origin (Swedish donors). The certified values are traceable to an internationally agreed reference measurement procedure (WHO's manual tilt tube technique) and the reference thromboplastin WHO IRP 67/40, through RBT/90 [10-12]. The procedures used to assign values are described in several publications and documents [13,14].

- PT (INR) calibrators from the Danish Institute for External Quality Assurance for Hospital Laboratories (DEKS) were used to get a link to the Danish PT (INR) level. The calibration materials from DEKS are freshly frozen pooled citrate-plasmas, which serve as national reference plasmas in Denmark. The DEKS calibration is a three point's calibration with a normal, therapeutic and high PT (INR). The assigned values come from three Nordic expert laboratories. The calibrators were analysed as samples on the comparison method at the start, in the middle and in the end of the evaluation.

  The calibrating systems from Equalis and DEKS are different with respect to the production of the materials as well as to the way the PT (INR)-values are set.

- At different occasions during the evaluation period, PT (INR) controls from Noklus were analysed on the comparison method.

5.4. The evaluation

5.4.1. *Planning of the evaluation*

Inquiry about an evaluation

iLine Microsystems S.L. via marketing manager Clara Grijelmo, applied to SKUP in November 2014 for an evaluation of the microINR system.

Protocol, arrangements and contract

In March 2015, the protocol for the evaluation was approved, and iLine Microsystems S.L and SKUP signed a contract for the evaluation. The Department of Medical Biochemistry in St. Olav Hospital, Trondheim, agreed to do the practical work with the evaluation under optimal conditions. Two PHCCs, Persaunet Legesenter and Hallset Legesenter from Sør-Trøndelag county, agreed to represent the intended users in this evaluation.

Training

Javier Alvariño from iLine Microsystems S.L, demonstrated the microINR system in the hospital laboratory. Hege Anette Martinsen (Orion Diagnostica) demonstrated the microINR system at PHCC1 and Britt S. Fredriksen (Orion Diagnostica) demonstrated the microINR system at PHCC2. All the evaluators were instructed to use mode A (see chapter 5.2) for sample application. The training reflected the training usually given to the intended users. The requesting company and the supplier were not allowed to contact or supervise the evaluators during the evaluation period.

5.4.2. *Evaluation sites*

The practical work was carried out during 17 weeks at the hospital laboratory and six weeks at the PHCCs, ending in August 2015.
The laboratory at the St. Olav university hospital has approximately 100 employees.
PHCC1 has three physicians, three health secretaries and one BLS. They use venous blood samples in their routine method for measurements of PT (INR). PHCC2 has five physicians, two health secretaries and one medical secretary. They use capillary blood samples in their routine method for measurements of PT (INR).

5.4.3. *The evaluation procedure under optimal conditions*

Internal analytical quality control
Internal analytical quality control for the microINR system (microINR Easy Control), one level, was measured each evaluation day.

Recruitment of patients
Patients who participated in this study were patients coming to the outpatient clinic for routine PT (INR) monitoring. Blood samples were collected from patients who had been stable on vitamin K antagonist treatment for a minimum of 4 weeks. Patients with known antiphospholipid syndrome (APS) were not included. Participation was voluntarily and verbal consent was considered sufficient based on national regulations.

Handling of the samples and measurements
All samples for measurement with the microINR system were capillary samples. The skin was disinfected with alcohol pads and the area dried completely before finger pricking. The samples were measured in duplicate using two skin-pricks from two separate fingers. Disposable lancing devices with depth settings 1,8 mm were used. In this evaluation, the second drop of capillary blood was measured after wiping off the first with a clean dry tissue/gauze. The sample was applied to the chip by approaching the finger to the meter; i.e. the meter was not moved. Three lot numbers of test chips were used in the evaluation. If error codes occurred the test was repeated, if possible, until a result was obtained.

Samples for the comparison method were obtained from venous puncture and collected into vacutainer tubes with 3,2% sodium citrate. The citrate samples were taken immediately before testing of the capillary samples on the microINR system. The tubes were inverted 8–10 times to ensure thorough mixing of the blood with the sodium citrate, and then underwent centrifugation for 10 minutes at 2200 g within two hours from sampling. Citrated fresh plasma was used for duplicate measurements of PT (INR) on the comparison method.

5.4.4. *The evaluation procedure for intended users*

Internal analytical quality control
Internal analytical quality control for the microINR system (microINR Easy Control), one level, was measured each evaluation day.

Recruitment of patients
Patients who participated in this study were patients at the PHCCs coming for routine PT (INR) monitoring. Blood samples were collected from patients who had been stable on vitamin K antagonist treatment for a minimum of 4 weeks. Patients with known antiphospholipid syndrome

(APS) were not included. Participation was voluntarily and verbal consent was considered sufficient based on national regulations.

Handling of the samples and measurements
All samples for measurements with the microINR system were capillary samples. The skin was disinfected with alcohol pads and the area dried completely before finger pricking. The samples were measured in duplicate using two skin-pricks from two separate fingers. Disposable lancing devices with depth settings 1,8 mm were used. In this evaluation, the second drop of capillary blood was measured after wiping off the first with a clean dry tissue/gauze. The sample was applied to the chip by approaching the finger to the meter; i.e. the meter was not moved. Three lot numbers of test chips were used in the evaluation. If error codes occurred the test was repeated, if possible, until a result was obtained.

Samples for the comparison method were obtained from venous puncture and collected into vacutainer tubes with 3,2% sodium citrate. The citrate samples were taken immediately before testing of the capillary samples on the microINR system. The tubes were inverted 8–10 times to ensure thorough mixing of the blood with the sodium citrate. The sample tubes were transported to the Department of Medical Biochemistry according to normal routine procedures at the PHCCs. The citrate plasma was analysed in duplicate for PT (INR) on the comparison method within 48 hours after sampling.

6. Results and discussion

Statistical expressions and calculations used by SKUP are shown in attachment 5.

6.1. Number of samples

Scheduled number of samples in this evaluation were 100 patient samples in duplicate under optimal conditions and 80 patient samples in duplicate measured by intended users. The hospital recruited 98 patients (SKUP ID 1-100). In the evaluation performed by intended users, PHCC1 and PHCC2 recruited 40 and 48 patients respectively (SKUP ID 111-150 and SKUP ID 201-248).

As a total, the results were spread over a wide range, but still there were few low and high results. Most of the results were within the interval 2,0 – 3,5 INR, consequently the results achieved in the hospital laboratory were divided into two, instead of three PT (INR) levels. This also provides an easier comparison with the results achieved in the two PHCCs, were the results usually are divided in only two levels because of the lower number of results.

An account of the number of samples not included in the calculations, is given below.

Missing results

- ID 112 (E18) and ID 113 (E05); only single measurements from microINR due to error codes. The single values were included in the calculation of bias and the assessment of accuracy.
- ID 122 (E03/E05), ID 125 (E03/E18/E17) and ID 149 (E05/E18); both results missing from microINR due to error codes.
- ID 46 and ID 47; the ID numbers were not used.
- ID 146, ID 148 and ID 150, only single measurements on the comparison method. The single values from the comparison method were still included in the calculations of bias and in the assessment of accuracy.
- ID 211-217, no results from the comparison method because the evaluator placed the venous samples in the fridge. These results were not included in the calculation of bias and the assessment of accuracy, but the results from microINR were included in the calculation of repeatability.
- The hospital laboratory did not analyse the internal analytical quality control at microINR eight of the days of the evaluation. In the PHCCs internal analytical quality control result from one day was missing. The results from the patient samples these days were still included in all the calculations.

Omitted results

- ID 119; the results from microINR were not included in the calculation of repeatability due to the use of two lot numbers, i.e. not identical conditions. The venous sample was not analysed on the comparison method within 48 hours as described in the evaluation procedure. The results from this patient were not included in any calculations.
- ID 121, ID 136 and ID 148; the second measurements from microINR were omitted from all the calculations due to deviation from the sampling procedure. The first measurements were included in the calculation of bias and the assessment of accuracy.

Excluded results

Statistical outliers in SKUP evaluations are detected by criterion promoted by Burnett [15].
- ID 9, ID 24 and ID 64; the results from the comparison method were classified as outliers according to Burnetts's model in the calculation of repeatability of the comparison method. The comparison method had good precision, and the statistical outliers were mostly a consequence of the low CV rather than actual differences between the duplicate measurements. The results were not included in the calculation of bias and the assessment of accuracy, but the results from microINR were included in the calculation of repeatability.

Failed measurements

Under optimal conditions, microINR reported the following error messages:
- 3 x E01; Meter cannot read the data matrix
- 1 x E03; Time out of the sample application countdown of 80 seconds.
- 1 x E05; Insufficient sample volume or not properly applied
- 2 x E17; Chip reading failure during testing time

Handled by intended users, microINR reported the following error messages:
- 2 x E03; Time out of the sample application countdown of 80 seconds
- 3 x E05; Insufficient sample volume or not properly applied
- 3 x E17; Chip reading failure during testing time
- 4 x E18; Sample mishandling or haematocrit out of range

Most of the error codes were related to the handling of the sample (11/372) = 3,0% errors
Eight error messages, E01 and E17, were interpreted as «technical errors», and six of these led to wasted microINR chips. The fraction of test wasted due to technical errors was estimated to: (6/372) x 100 = 1,6%.

The SKUP recommendation of an incident of test wasted due to technical errors ≤2,0% was achieved.

6.5. Evaluation of user-friendliness

6.5.1. *Questionnaire to the evaluators*

The most important response regarding user-friendliness comes from the intended users themselves. The end-users often emphasize other aspects than those pointed out by more extensively trained laboratory personnel.

At the end of the evaluation period, the intended user fills in a questionnaire about the user-friendliness of the instrument. SKUP has prepared detailed instructions for this.

The questionnaire is divided into four subareas:

Table A) Rating of the information in the manual / insert / quick guide
Table B) Rating of operation facilities. Is the system easy to handle?
Table C) Rating of time factors for the preparation and the measurement
Table D) Rating of performing internal and external analytical quality control

The intended users fills in table A and B. SKUP fills in table C and D and in addition, topics marked with grey colour in table A and B.

In the tables, the first column shows what is up for consideration. The second column in table A and B shows the rating by the individual users at the evaluation sites. The last three columns show the rating options. The overall ratings from all the evaluating sites are marked in coloured and bold text. The last row in each table summarises the total rating in the table. The total rating is an overall assessment by SKUP of the described property, and not necessarily the arithmetic mean of the rating in the rows. Consequently, a single poor rating can justify an overall poor rating, if this property seriously influences on the user-friendliness of the system.

Unsatisfactory and intermediate ratings are marked with a number and explained below the tables. The intermediate category covers neutral ratings assessed as neither good nor bad.

An assessment of the user-friendliness is subjective, and the topics in the questionnaire may be emphasised differently by different users. The assessment can therefore vary between different persons and between the countries. This will be discussed and taken into account in the overall assessment of the user-friendliness.

Comment

In this evaluation, the user-friendliness was assessed by PHCC1 (the opinion of one BLS and three health secretaries) and PHCC2 (the opinion of two health secretaries and one medical secretary). PHCC2 did not evaluate table A.

Table A. Rating of the information in the manual

Topic	Rating	Assessment	Assessment	Assessment
General impression	I[1]	Satisfactory	**Intermediate**	Unsatisfactory
Table of contents	S	**Satisfactory**	Intermediate	Unsatisfactory
Preparations / Pre-analytic procedure	S	**Satisfactory**	Intermediate	Unsatisfactory
Specimen collection	I[2]	Satisfactory	**Intermediate**	Unsatisfactory
Measurement procedure	I[2]	Satisfactory	**Intermediate**	Unsatisfactory
Reading of result	S	**Satisfactory**	Intermediate	Unsatisfactory
Description of the sources of error	S	**Satisfactory**	Intermediate	Unsatisfactory
Help for troubleshooting	I[2]	Satisfactory	**Intermediate**	Unsatisfactory
Readability / Clarity of presentation	U[3]	Satisfactory	Intermediate	**Unsatisfactory**
Keyword index*		Satisfactory	Intermediate	Unsatisfactory
Measurement principle	I[4]	Satisfactory	**Intermediate**	Unsatisfactory
Available insert in Danish, Norwegian, Swedish	S	**Satisfactory**	Intermediate	Unsatisfactory
Total rating by SKUP			**Intermediate**	

*Not rated in this evaluation due to the small size of the manual
[1]In general there are too many difficult words in the manual
[2]Some of the instructions are ambiguous
[3]Small print
[4]Explanation of the measurement principal is affected by poor translation into Norwegian

Table B. Rating of operation facilities

Topic	Rating	Assessment	Assessment	Assessment
To prepare the test / instrument	I[1], U[1]	Satisfactory	**Intermediate**	**Unsatisfactory**
To prepare the sample	S, U[2]	**Satisfactory**	Intermediate	**Unsatisfactory**
Application of specimen	U[3], U[3]	Satisfactory	Intermediate	**Unsatisfactory**
Specimen volume	S, S	**Satisfactory**	Intermediate	Unsatisfactory
Number of procedure step	S, I[4]	**Satisfactory**	**Intermediate**	Unsatisfactory
Instrument / test design	I[5], U[6]	Satisfactory	**Intermediate**	**Unsatisfactory**
Reading of the test result	S, S	**Easy**	Intermediate	Difficult
Sources of errors	U[7], -	Satisfactory	Intermediate	**Unsatisfactory**
Cleaning / Maintenance	S, -	**Satisfactory**	Intermediate	Unsatisfactory
Hygiene, when using the test	S, U[8]	**Satisfactory**	Intermediate	**Unsatisfactory**
Size and weight of package	S, S	**Satisfactory**	Intermediate	Unsatisfactory
Storage conditions for tests, unopened package*	S	**+15 to +30°C**	+2 to +8°C	–20°C
Storage conditions for tests, opened package**		+15 to +30°C	+2 to +8°C	–20°C
Environmental aspects: waste handling	S	**No precautions**	Sorted waste	Special precautions
Intended users	S	**Health care personnel or patients**	Laboratory experience	Biomedical laboratory scientists
Total rating by SKUP				**Unsatisfactory**

*Storage temperature for the microINR chips is +2 to +25°C

**Not rated in this evaluation due to the single-pack concept of the microINR chip

[1]The period of time from the meter is switched on to the blood can be placed on the test chip, is too long (comment from SKUP: iLine Microsystems S.L. informs that the start-up procedure takes from 50-70 seconds under standard conditions.)

[2]The meter could not be moved (comment from SKUP: as instructed by iLine Microsystems S.L. and Orion Diagnostica during the training, see 5.2 and 5.4.1) and the finger could not touch the test chip when applying the sample. The system is too sensitive and has too many sources of error.

[3]The test chip is too sensitive to the way the sample is applied

[4]Too many details to take care of when the sample is to be placed on the microINR chip
[5]Good-sized meter and clear display, but difficult to analyse the internal quality control material
[6]When the test chip is inserted in the meter and the sample is about to be placed on the chip, it becomes too little space to do it properly. It would be better if the meter could be lifted towards the finger
[7]Many error codes due to the sensitivity to the way the sample is applied
[8]It is possible to spill blood when a used test chip is removed from the meter

Additional negative comment: The battery needs to be recharged more often than specified in the manual.

Table C. Rating of time factors (filled in by SKUP)

Topic	**Assessment**	**Assessment**	**Assessment**
Required training time	**<2 hours**	2 to 8 hours	>8 hours
Durations of preparations / Pre-analytical time	**<6 min.**	6 to 10 min.	>10 min.
Duration of analysis	**<10 min.**	10 to 20 min.	>20 min.
Stability of test, unopened package	**>5 months**	3 to 5 months	<3 months
Stability of test, opened package*	>30 days	14 to 30 days	<14 days
Stability of quality control material, unopened**	**>5 months**	3 to 5 months	<3 months
Stability of quality control material, opened***	>6 days or disposable	2 to 6 days	≤1 day
Total rating by SKUP	**Satisfactory**		

*Not rated in this evaluation due to the single-pack concept of the microINR chip
**The stability of the internal analytical control material microINR Easy Control is >5 months if the control material is stored at +2 to +8°C
***Not rated in this evaluation due to the fact that clotting will start immediately after reconstitution

Additional negative comments: There is no information about the storage conditions of the chip in the manual, only in the package insert of the chip.

Table D. Rating of analytical quality control (filled in by SKUP)

Topic	Assessment	Assessment	Assessment
Reading of the internal quality control	**Satisfactory**	Intermediate	Unsatisfactory
Usefulness of the internal quality control	Satisfactory	**Intermediate**[1]	Unsatisfactory
External quality control	**Satisfactory**	Intermediate	Unsatisfactory
Total rating by SKUP	**Satisfactory**		

[1]The imprecision achieved with the internal analytical control material microINR Easy Control is above the imprecision of the patient samples, i.e. less possibilities to discover errors in the analytical system

7. References

1. Christensen N.G., Monsen G. & Sandberg S. Utprøving av analyseinstrumenter, 1997. Alma Mater Forlag.
2. ISO/FDIS 17593:2007, Clinical laboratory testing and in vitro medical devices – Requirements for in vitro monitoring systems for self-testing of oral anticoagulant therapy.
3. Hyltoft Petersen P. *et al.* Combination of analytical quality specifications based on biological within- and between-subject variation. Ann Clin Biochem 2002; **39** (6): 543 – 550.
4. Fraser C.G. & Hyltoft Petersen P. Quality goals in external quality assessment are best based on biology. Scand J Clin Lab Invest 1993; **53** suppl 212. Chapter I. Quality planning.
5. Ricos C. *et al.* Current databases on biological variation: pros, cons and progress. Scand J Clin Lab Invest 1999; **66** (4): 337 – 349.
6. Kjeldsen J., Lassen J.F., Hyloft Petersen P. & Brandslund I. Biological variation of international normalized ratio for prothrombin times, and consequences in monitoring oral anticoagulant therapy: computer simulation of serial measurements with goalsetting for analytical quality. Clin Chem 1997; **43** (11): 2175 – 2182.
7. Kvalitetskrav og kvalitetsvurdering for hyppigt udførte klinisk biokemiske og klinisk mikrobiologiske analyser i almen praksis. Konsensus dokument udarbejdet af Laboratorieudvalget under Sygesikringens og PLO´s Faglige Udvalg vedr. Almen Praksis i samarbejde med DEKS og Dansk Selskab for Klinisk Biokemi's Videnskabelige udvalg. Nov 2003.
   www.skup.nu (menu The SKUP evaluations – Quality goals).
8. Kvalitetssikring og kvalitetskrav til laboratoriemedicinske aktiviteter i almen praksis. Udarbejdet af Regionernes Lønnings- og Takstnævn (RTLN) og Praktiserende Lægers Organisation (PLO). 2010.www.skup.nu (menu The SKUP evaluation – Quality goals).
9. http://www.ifcc.org/ifcc-scientific-division/sd-committees/c-npu/npusearch/
10. Van den Besselaar A.M. Multicentre study of replacement of the international reference preparation for thromboplastin rabbit plain. Thromb Haemost 1993; **70**: 794 – 799.
11. Van den Besselaar A.M. & Houdijk W.P. Use of lyophilized calibrant plasmas for simplified international normalized ratio determination with a human tissue factor thromboplastin reagent derived from cultured human cells. Clin Chem 2003; **49** (12): 2006 – 2011.
12. WHO. Expert committee on biological standardization − Requirements for thromboplastins and plasma used to control oral anticoagulant therapy. World Health Organization, Geneva Technical Report Series 889, 48th Report 1999.
13. Hillarp A. *et al*. Local INR calibration of the Owren type prothrombin assay greatly improves the intra- and interlaboratory variation. Thromb Haemost 2004; **91**: 3300 – 3307.
14. Arbetsbeskrivning A147, ver 1.0, 2013. Arbetsbeskrivning för Equalis och Expertgrupp vid åsättande av INR-värden till kalibratorer och kontrollmaterial för bestämning av protrombinkomplex enligt Owren. Equalis, Uppsala.
15. Burnett RW. Accurate estimation of standard deviations for quantitative methods used in clinical chemistry. Clinical Chemistry 1975; **21** (13): 1935 – 1938.

Statistical expressions and calculations

This chapter with standardised text deals with the statistical expressions and calculations used by SKUP. The statistical calculations will change according to the type of evaluation. The descriptions in this document are valid for evaluations of quantitative methods with results on the ratio scale.

Statistical terms and expressions

The definitions in this section come from the ISO/IEC Guide 99; International Vocabulary of Metrology, VIM [a].

Precision

Definition: Precision is the closeness of agreement between measured quantity values obtained by replicate measurements on the same or similar objects under stated specified conditions.

Precision is measured as *imprecision*. Precision is descriptive in general terms (good, poor e.g.), whereas the imprecision is expressed by means of the standard deviation (SD) or coefficient of variation (CV). SD is reported in the same PT (INR) as the analytical result. CV is usually reported in percent.

To be able to interpret an assessment of precision, the precision conditions must be defined.
Repeatability is the precision of consecutive measurements of the same component carried out under identical measuring conditions (within the measuring series).
Reproducibility is the precision of discontinuous measurements of the same component carried out under changing measuring conditions over time.

Trueness

Definition: Trueness is the closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value.

Trueness is inversely related to systematic measurement error. Trueness is measured as *bias*. Trueness is descriptive in general terms (good, poor e.g.), whereas the bias is reported in the same PT (INR) as the analytical result or in percent.

Accuracy

Definition: Accuracy is the closeness of agreement between a measured quantity value and the true quantity value of a measurand.

Accuracy is not a quantity and cannot be expressed numerically. A measurement is said to be more accurate when it offers a smaller measurement error. Accuracy can be illustrated in a difference-plot. Accuracy is descriptive in general terms (good, poor e.g.).

a. ISO/IEC Guide 99:2007, International vocabulary of metrology – Basic and general concepts and associated terms, VIM, 3rd edition, JCGM 200:2008.

Statistical calculations

Statistical outliers

The criterion promoted by Burnett [b] is used for the detection of outliers. The model takes into consideration the number of observations together with the statistical significance level for the test. The significance level is set to 5%. The segregation of outliers is made with repeated truncations, and all results are checked. Where the results are classified according to different concentration levels, the outlier-testing is carried out at each level separately. Statistical outliers are excluded from the calculations.

Calculation of imprecision

The precision of the evaluated method is assessed by use of paired measurements of genuine patient sample material. The results are usually divided into three concentration levels, and the estimate of imprecision is calculated for each level separately, using the following formula [c,d]:

$$SD = \sqrt{\frac{\sum d^2}{2n}}$$

d = difference between two paired measurements
n = number of differences

This formula is used when the standard deviation can be assumed reasonable constant across the concentration interval. If the coefficient of variation is more constant across the concentration interval, the following formula is preferred:

$$CV = \sqrt{\frac{\sum (d/m)^2}{2n}}$$

m = mean of paired measurements (formula 2)

The two formulas are based on the differences between paired measurements. The calculated standard deviation or CV is still a measure of the imprecision of single values. The imposed condition for using the formulas is that there is no systematic difference between the 1st and the 2nd measurement of the pairs. The CV is given with a 90% confidence interval.

Calculation of bias

The mean deviation (bias) at different concentration levels is calculated based on results achieved under optimal measuring conditions. A paired t-test is used with the mean values of the duplicate results on the comparison method and the mean values of the duplicate results on the evaluated method. The mean difference is shown with a 95% confidence interval.

Assessment of accuracy

The agreement between the evaluated method and the comparison method is illustrated in a difference-plot. The x-axis represents the mean value of the duplicate results on the comparison method. The y-axis shows the difference between the first measurement on the evaluated method and the mean value of the duplicate results on the comparison method. The number of results within the quality goal limits is counted and assessed.

b. Burnett RW. Accurate estimation of standard deviations for quantitative methods used in clinical chemistry. Clinical Chemistry 1975; **21** (13): 1935 – 1938.
c. Saunders E. Tietz textbook of clinical chemistry and molecular diagnostics, 2006. Chapter 14, Linnet K., Boyd J. Selection and analytical evaluation of methods – with statistical techniques. Elsevier Saunders ISBN 0-7216-0189-8.
d. Fraser C.G. Biological variation: From principles to practice, 2006. Chapter 1, The Nature of Biological Variation. AACC Press ISBN 1-890883-49-2.

Results

Table 4. Repeatability, PT (INR) venous citrate samples, the comparison method

PT (INR) level Comparison method	n	Excluded results	Mean value (interval), PT (INR)	CV (90% CI) %
<2,5	38	1*	2,1 (1,05 – 2,49)	1,1 (0,9 – 1,3)
≥2,5	57	2*	3,0 (2,52 – 4,67)	1,0 (0,8 – 1,2)

* The given numbers of results (n) are counted before the exclusion of results. Mean and CV are calculated after the exclusion of results. ID 9, ID 24 and ID 64 are statistical outliers according to Burnett's model [15] and therefore excluded. An account of the number of samples is given in section 6.1.

Table 5. Equalis PT (INR) calibrators measured on the comparison method

Material	Certified value, PT (INR) (uncertainty)	Date	n	Mean value, PT (INR) instrument 1	Mean value, PT (INR) instrument 2
Equalis INR calibrator Low	**1,05** (0,96 – 1,14)	14.04.15	5	1,07	1,04
		27.07.15	5	1,07	1,08
Equalis INR calibrator High	**3,14** (2,57 – 3,71)	14.04.15	5	3,09	3,12
		27.07.15	5	3,20	3,22
Equalis INR control	**2,48** (2,09 – 2,87)	14.04.15	5	2,40	2,43
		27.07.15	5	2,49	2,49

Table 6. DEKS PT (INR) calibrators measured on the comparison method

Material	Assigned value, PT (INR) (uncertainty)	Date	n	Mean value, PT (INR) instrument 1	Mean value, PT (INR) instrument 2
DEKS INR calibrator Normal	**1,00** (0,98 – 1,03)	14.04.15	5	1,02	0,99
		13.05.15	5	0,99	1,01
		27.07.15	5	1,02	1,04
DEKS INR calibrator Therapeutic	**2,26** (2,19 – 2,33)	14.04.15	5	2,23	2,22
		13.05.15	5	2,24	2,24
		27.07.15	5	2,25	2,27
DEKS INR calibrator High	**3,74** (3,59 – 3,89)	14.04.15	5	3,81	3,86
		13.05.15	5	3,83	3,82
		27.07.15	5	3,89	3,91

Table 7. Repeatability, PT (INR) capillary samples microINR. Results achieved under optimal conditions

PT (INR) level microINR	n	Excluded results	Mean value (interval) PT (INR)	CV (90% CI) %
<2,5	34	0	2,0 (1,00 – 2,45)	*
≥2,5	64	0	3,0 (2,50 – 4,70)	6,0 (5,3 – 7,0)

An account of the number of samples is given in section 6.1.
* If the repeatability is calculated without taking into consideration the systematic difference pointed out between the paired measurements in this level, the CV is 5,6. If the average systematic difference is removed from all paired results in this level, the estimated CV becomes 4,1.

Table 8. Bias, PT (INR) capillary samples microINR. Results achieved under optimal conditions

PT (INR) level Comparison method	n	Excluded results	Mean value Comparison method, PT (INR)	Mean value PT (INR)	Bias (95% CI), INR	Bias , %
<2,5	38	1*	2,1	2,1	+0,06 (0,002 – +0,13)	3,1
≥2,5	57	2*	3,0	3,0	-0,04 (-0,11 – +0,02)	-1,4

* The given numbers of results (n) are counted before the exclusion of results. Mean and bias are calculated after the exclusion of results. ID 9, ID 24 and ID 64 are statistical outliers according to Burnett's model [15] and therefore excluded. An account of the number of samples is given in section 6.1.

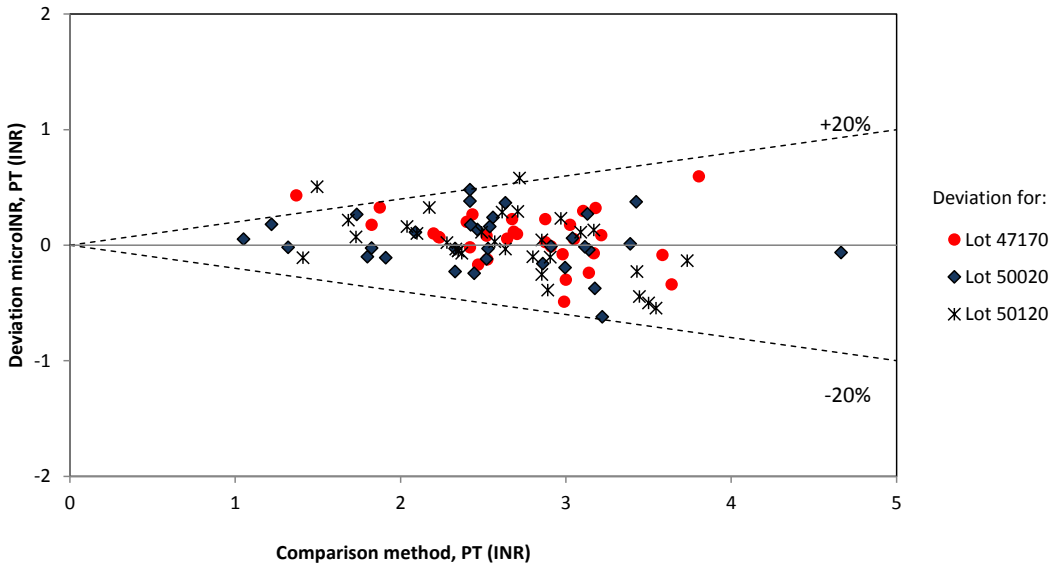

Figure 2. Accuracy of PT (INR) on the microINR system under optimal conditions. The x-axis represents the mean PT (INR) result of the corresponding sample of the comparison method. The y-axis represents the deviation in PT (INR) of the first capillary sample measurement on microINR from the mean result of the comparison method. Different lots are illustrated as Lot 47170 (●), lot 50020 (◆) and lot 50120 (✳). Stippled lines represent allowable deviation limits of ±20%. Number of results (n) = 95. An account of the number of samples, and excluded and missing results, is given in section 6.1.

Table 9. Repeatability, PT (INR) capillary samples microINR. Results achieved by intended users

PT (INR) level microINR	n	Excluded results	Mean value (interval) PT (INR)	CV (90% CI) %
PHCC1				
< 2,5	13	0	2,2 (1,85 – 2,45)	5,0 (3,7 – 7,8)
≥2,5	18	0	2,9 (2,50 – 3,65)	6,1 (4,8 – 8,6)
PHCC2				
< 2,5	24	0	2,1 (1,35 – 2,45)	*
≥2,5	24	0	3,1 (2,50 – 7,65)	6,3 (5,1 – 8,3)

An account of the number of samples is given in section 6.1.
* If the repeatability is calculated without taking into consideration the systematic difference pointed out between the paired measurements in this level from PHCC2, the CV is 4,7. If the average systematic difference is removed from all paired results in this level, the estimated CV becomes 4,2.

Table 10. Bias, PT (INR) capillary samples microINR. Results achieved by intended users

PT (INR) level Comparison method	n	Excluded results	Mean value Comparison method, PT (INR)	Mean value PT (INR)	Bias (95% CI), INR	Bias, %
PHCC1						
< 2,5	19	0	2,18	2,26	+0,08 (-0,03 – +0,20)	3,8
≥2,5	17	0	3,06	2,94	-0,11 (-0,25 – +0,02)	-3,8
PHCC2						
< 2,5	22	0	2,12	2,15	+0,03 (-0,05 – +0,11)	1,5
≥2,5	19	0	3,24	3,16	-0,08 (-0,25 – +0,09)	-2,6

An account of the number of samples is given in section 6.1.

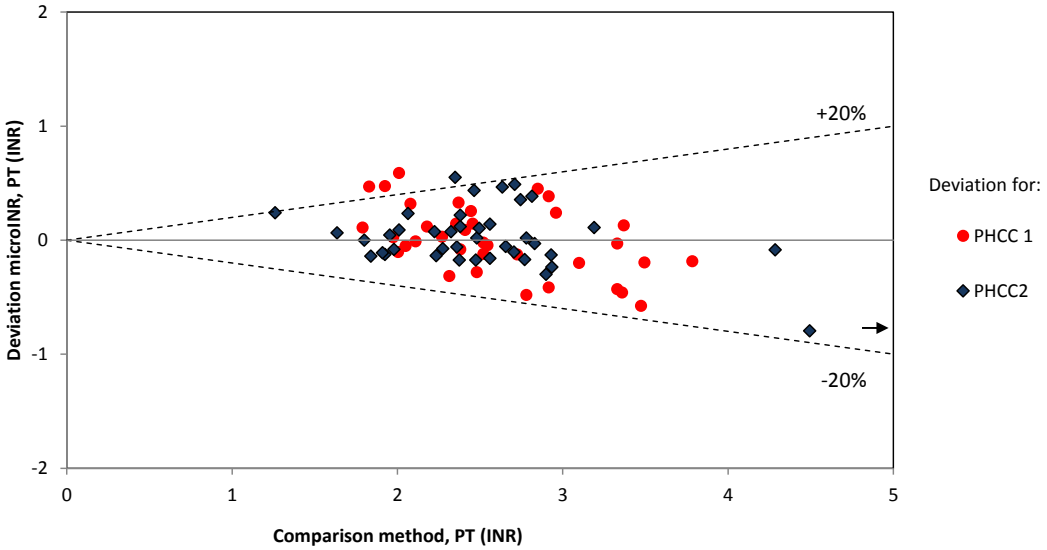

Figure 3. Accuracy of PT (INR) on the microINR system achieved by intended users. The x-axis represents the mean PT (INR) result of the corresponding sample of the comparison method. The y-axis represents the deviation in PT (INR) of the first capillary samples measurement on microINR from the mean result of the comparison method. The results from PHCC1 are represented with the symbol (●) and results from PHCC2 with the symbol (◆). Stippled lines represent allowable deviation limits of ±20%. Number of results (n) = 77. One result at 8.5 INR, is not shown in the figure, but illustrated as → .The result was within the allowable deviation limits. An account of the number of samples, and excluded and missing results, is given in section 6.1.