Nova StatStrip Glucose and ß-Ketone Hospital Meter System

A system for measurement of
the concentration of glucose and ß-ketone
manufactured by
Nova Biomedical Corporation, USA

A report from an evaluation of glucose measurements organised by SKUP

Evaluated at the request of
the distributor for the Nordic countries
A. Menarini Diagnostics Nordic Countries
in Sweden

SKUP/2013/85

The report is written by SKUP in the spring 2012 and after that changed in details
Main author is Arne Mårtensson
SKUP in Sweden

Table of contents

Raw data is sent electronically to the company requesting the report.

1 Summary

Background
The Nova StatStrip™ Glucose and ß-Ketone Hospital Meter System (StatStrip) is intended for measuring glucose in fresh capillary, venous, or arterial blood including neonate blood, by health care professionals. The glucose results are calibrated to agree with a method measuring the concentration in plasma.

StatStrip is manufactured by Nova Biomedical Corporation, USA. The agent for the system in the Nordic countries is A. Menarini Diagnostics Nordic Countries in Sweden, who also requested this evaluation. The evaluation was performed in the period February to November 2011.

The StatStrip system consists of the StatStrip meter and the StatStrip test strips. The sample volume, 1,2 µL, is aspirated to the test strip by capillary draw.

The aim of this study was to evaluate
- the analytical quality of glucose measurements
- the user-friendliness

The evaluation was performed both under optimal conditions when operated by an experienced biomedical laboratory scientist and under "real life" conditions when operated by the intended users, nurses and midwifes

Materials and methods
StatStrip is evaluated under different conditions and with tighter accuracy goals compared to SKUP evaluations of meters for glucose self-monitoring.

In a hospital laboratory, an experienced biomedical laboratory scientist carried out StatStrip measurements on arterial samples. In one hospital ward, the measurements were carried out by nurses on capillary samples from adult persons with diabetes and in another ward midwifes measured on venous samples from healthy newborn children. Three lots of test strips were used.

The comparison method was the routine method for P—Glucose in the Karolinska University Laboratory, Huddinge. The method is accredited. It is the Roche hexokinase method, Gluco-quant Glukos/HK, applied on a Modular Analytics P instrument from Roche Diagnostics.

The analytical quality goals set by SKUP for this evaluation were that repeatability should not exceed 4% CV and that at least 95% of the results should fall within ±0,83 mmol/L at glucose concentrations <5,6 mmol/L and within ±15% at glucose concentrations ≥5,6 mmol/L, from the comparison method results.

Results
In the hospital laboratory. With arterial samples the obtained repeatability was 3% CV. Patient sample results showed a bias of approximately +0,2 mmol/L. Ninety-nine percent of the results were inside the accuracy goal limits. The StatStrip results were not influenced by haematocrit, pO_2, pH and sodium concentrations in the samples within the examined intervals (for haematocrit 20 to 47%). The results in the hospital laboratory fulfilled the quality goals.

At the hospital wards. With venous and capillary samples the obtained repeatability was from 5 to 7% CV. The quality goal for imprecision was not fulfilled. Patient sample results showed a bias of about +0,3 mmol/L. The accuracy goal was fulfilled in the hospital ward using venous samples, with 95% of the results inside the limits, but not in the ward using capillary samples, where 93% of the results were inside the limits.

See an overview of the analytical quality results on page 32.

User-friendliness. The evaluators' general opinion was that StatStrip was user-friendly and easy to handle. For most of the items StatStrip got the best assessment "Satisfactory".
The evaluators remarked that the operation of StatStrip requires several procedure steps; scanning the test strip lot number, operator ID and patient ID.

Conclusion

In the hands of an experienced biomedical laboratory scientist, with arterial samples from adult intensive care patients, the analytical quality of StatStrip was good and fulfilled the quality goals. When nurses and midwifes measured venous and capillary patient samples, the quality goal for imprecision was not fulfilled. The quality goal for accuracy was fulfilled in one ward measuring venous neonatal samples but not in another ward measuring capillary adult samples. The accuracy of the StatStrip results was not influenced by haematocrit, pO_2, pH and sodium. StatStrip was easy to handle.

Comments from the manufacturer or requesting company

There are no comments from Nova Biomedical or from A. Menarini Diagnostics Nordic Countries.

2 Abbreviations

ADA	American Diabetes Association
CI	Confidence Interval
C-NPU	Committee on Nomenclature, Properties and Units
CV	Coefficient of Variation
DAK-E	Danish Quality Unit of General Practice
EQA	External Quality Assessment
Equalis	External quality assurance in laboratory medicine in Sweden
HK	Hexokinase
IFCC	The International Federation of Clinical Chemistry and Laboratory Medicine
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
Menarini	A. Menarini Diagnostics Nordic Countries in Sweden
NIST	National Institute of Standards & Technology
NOKLUS	Norwegian Quality Improvement of Primary Care Laboratories
SKUP	Scandinavian evaluation of laboratory equipment for primary health care
SRM	Standard Reference Material
StatStrip	Nova StatStrip™ Glucose and ß-Ketone Hospital Meter System

3 Quality goals

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.

3.1 Analytical quality goals for P—Glucose tests

Various ways of setting analytical quality goals for P—Glucose determinations are presented below. Ideally analytical quality goals should be set according to medical demands.
For P—Glucose it is natural that the goals are set differently depending on the intended use of the measurement results. StatStrip is intended for the measurement of P—Glucose in fresh capillary, venous, arterial, and neonate whole blood in hospital care and the quality goals should be set accordingly.

3.1.1 Analytical quality goals based on recommendations from professionals/experts

For glucose instruments intended for monitoring, it applies to produce values with good precision [1]. According to ADA the imprecision should be less than 5% CV [2]. Other authors also recommend the same imprecision requirements [3].

Analytical quality goals for accuracy when monitoring P—Glucose is also recommended in ”Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus” [4]. ISO 15197:2003 claims the following minimum requirements:

Ninety-five percent (95%) of measurements performed by health personnel should fall within ±0,83 mmol/L at glucose concentrations <4,2 mmol/L and within ±20% at glucose concentrations ≥4,2 mmol/L when compared with an established glucose method.

A revision of ISO 15197 is going on in June 2012 and it is likely that the tolerance limits will be lowered from ±20% to ±15%:

Ninety-five percent (95%) of measurements performed by persons with diabetes should fall within ±0,83 mmol/L at glucose concentrations <5,6 mmol/L and within ±15% at glucose concentrations ≥5,6 mmol/L when compared with an established glucose method.

In the NICE-SUGAR Study [5] the main conclusion was that that tight glucose control to normal glucose levels increased mortality among adults in the intensive-care units compared to conventional glucose control. In a following discussion about measurement quality between Mahoney and Cembrowski, the latter argued for maximal total error of 12,5% [6]. Cembrowski also referred to a recent FDA meeting held 17 March 2010 where the same maximal total error limit had been proposed [7].

Boyd [8] has specified quality requirements for glucose meters from a simulation study where the effect on errors in insulin dose were studied. They found that large errors of insulin dose (two-step or greater) occurred >5% of the time when the CV and/or bias exceeded 10 – 15%.

The Laboratory Committee under the Professional Committee in Denmark has in November 2003 recommended the analytical quality requirements for venous and capillary blood. The specification in Table 1 is valid for sample collection after 15 minutes rest in sitting position.

Table 1. Danish national quality requirements

	Hospital laboratory	Primary health care
Bias	<1,5% (venous plasma)	<3% (capillary whole blood)
CV	<2,5% (venous plasma)	<4% (capillary whole blood)

The requirements for capillary blood correspond to a ±10% allowable deviation for the single results.

Matrix effects: Some of the near patient equipment for measurements of glucose can only be used with fresh capillary whole blood. In external quality schemes the matrix of the used sample materials cause false results, which not necessarily show the same deviations as the patient samples.

In Norway NOKLUS in 2008 suggested a quality goal for glucose instruments for use in primary care centres and nursing homes in Norway [9]. The quality goal is that they should show a total error (the sum of imprecision and bias) ≤10%.

In Sweden the analytical quality goals for P—Glucose are agreed upon in a national consensus document [10] from 2007: The agreement gives sanctions to the analytical quality goals set differently depending on the intended use of the measurement results. Methods for diabetes monitoring in health care should below 4,2 mmol/L deviate less than ±0,83 and at 4,2 mmol/L and above less than ±20%. It is enough to compare the results with a selected comparison method.

The intended use of the StatStrip system, as declared by Nova Biomedical Corporation (Nova Biomedical), is for monitoring of glycemia in hospitalised patients. However, after discussions with representatives for Nova Biomedical it was decided to evaluate the meter against tighter quality goals than ±20%.

In several procurements in Sweden there is a wish to assess glucose meters against really tight limits. The idea is not that all results have to fulfil the limits but to make it possible to grade the analytical quality of the best meters by presenting the achieved percentage of results within certain limits. The following requirements have then been proposed:

Give the percentage of measurement results within ±0,40 mmol/L at glucose concentrations <4,0 mmol/L and within ±10% at glucose concentrations ≥4,0 mmol/L when compared with an established hexokinase method for plasma glucose.

To help the Swedish purchasers SKUP has decided to calculate the achieved percentage with these limits also in the SKUP reports. Note that in the SKUP evaluations of meters for glucose self monitoring there are usually no results below 4,0 mmol/L so the ±0,40 mmol/L limits are not applied in those reports.

3.2 Quality goals for user-friendliness

The evaluation of user-friendliness is carried out by asking the evaluating persons (end-users) to fill in a questionnaire. In the questionnaire the user-friendliness is divided into four sub-areas:

- Information in manual and insert
- Time factors during the measurement and preparation
- Performing internal and external quality control.
- Operation facilities. Is the system easy to handle?

Evaluation of user-friendliness is rated with the following scale:

”Green” stands for satisfactory
”Yellow” stands for less satisfactory
”Red” stands for unsatisfactory

3.3 Evaluation conditions

StatStrip is evaluated with stricter accuracy goals and under different conditions compared to evaluations of glucose meters for self-monitoring.

- The operators in this evaluation are less than ten health care professionals while in evaluations of glucose meters for self-monitoring they are around 90 persons with diabetes.
- The samples in this hospital laboratory evaluation derived from patients in the Intensive-Care Unit and can be presumed to contain normally occurring drugs and metabolites which might interfere with the measurements.
- The samples in the Maternity Ward evaluation derived from newborn children with high haematocrit which is a challenge for glucose meters.
- The concentration range covered in this evaluation is wider than in SKUP evaluations of glucose meters for self-monitoring.
- The SKUP procedures for capillary sample collection, in evaluations of glucose meters for self-monitoring, deviate from the procedures in the ISO standard 15197. The main difference is that ISO standard prescribes that the capillary sample collected from each patient is mixed in a tube before it is used as a single sample source both for replicate measurements (imprecision) and for the measurements used for comparison (accuracy). SKUP however collects and measures multiple capillary samples directly from finger punctures. The SKUP procedures are selected intentionally to evaluate the analytical quality as when capillary sampling is used in practice. In this evaluation the capillary samples at the Endocrine clinic were collected according to the SKUP procedures. The quality goals in the standard, draft EN ISO 15197:2013, are not fulfilled in this part of the evaluation. It is not known if the ISO quality goals had been fulfilled if the ISO procedures had been used instead of the SKUP procedures.
- For the hospital laboratory and the Maternity ward, samples were collected in tubes and mixed before the measurements according to procedures in the ISO standard.

3.4 SKUP's quality goals in this evaluation

Based on the discussion about alternative quality goals above, it was agreed in the protocol to assess the results from the evaluation of StatStrip against the following quality goals:

Repeatability (CV_a) ..... ≤4%

Allowable deviation
in the individual result from the comparison method result
for glucose concentrations <5,6 mmol/L ..... ≤±0,83 mmol/L
and for glucose concentrations ≥5,6 mmol/L ..... ≤±15%
Required percentage of individual results
within the above allowable deviations ..... ≥95%

Fraction of technical errors ..... ≤2%

User-friendliness ..... Satisfactory

4 Materials and methods

4.1 Definition of the measurand

The IFCC and the IUPAC work in a joint committee on nomenclature, properties and units (C-NPU). The descriptions of clinical laboratory tests are listed in the ”NPU database” [11]. In the database the recommended name is given for the measurand together with which unit the result should be reported in.

In this report the measurand is called P—Glucose and the results are expressed in the unit mmol/L. It is the measurand intended to be measured with the Nova StatStrip even if whole blood is the normal sample material for this system.

4.2 The evaluated measurement system: StatStrip

The information in this section is collected from Nova Biomedical.
The Nova StatStrip™ Glucose and ß-Ketone Hospital Meter System (StatStrip) consists of the instrument, the StatStrip meter and the disposable StatStrip test strips.

Figure 1. StatStrip meter with a test strip

4.2.1 The measurement principle of StatStrip

StatStrip is a modified glucoseoxidase-based amperometric test system with haematocrit and chemical interference corrections. The strength of the generated current at the electrode is proportional to the glucose concentration of the sample. The single-use StatStrip test strip is thus a biosensor.

Glucose + Enzymes (oxidized form) ——> Gluconic Acid + Enzymes (reduced form) Step 1

Enzymes (reduced form) + Ferricyanide —> Enzymes (oxidized form) + Ferrocyanide Step 2

$$\text{Ferrocyanide} \xrightarrow[\text{Electrode}]{e^-} \text{Ferricyanide} \qquad \text{Step 3}$$

On the StatStrip test strip there is more than one-measuring well. One of StatStrip’s measuring wells measures haematocrit to make it possible for the meter to correct the glucose

result for abnormal haematocrit levels. An additional interference measurement well is used to measure and eliminate electrochemical interferences from maltose, galactose, oxygen, and other electrochemical interferents.

4.2.2 Basic facts about StatStrip

Table 2. Basic specifications by the manufacturer

Sample material:	Whole blood: Capillary, venous, arterial, and neonate.
Sample volume:	1,2 μL
Measuring principle:	Glucoseoxidase-based amperometric test system
Measuring range:	0,6 to 33,3 mmol/L
Tolerated haematocrit interval	20 to 65%
Measurement duration:	6 seconds
Memory capacity of the instrument:	1000 patient measurements 200 quality control measurements 4000 operators
Electrical power supply:	Battery: Rechargeable Li-polymer 3,7 V, 2000 mAh Life: 6-8 hours in use (approximately 40 tests with barcode scans)/12-24 hours standby
Dimensions and weight:	Width: 82,5 mm Depth: 46 mm Height: 153 mm Weight: 360 g

This is an extract from more comprehensive tables in Attachment 1.

4.2.3 Analysing a patient sample with StatStrip

A guide from the StatStrip manual is reprinted in Attachment 1. The instructions in the guide were followed during the evaluation.

4.2.4 Intended use of StatStrip

According to the manufacturer, StatStrip is intended for in vitro diagnostic use by health care professionals and for point-of-care usage for the quantitative measurement in fresh capillary, venous, arterial and neonate whole blood samples. The glucose results are calibrated to agree with a method measuring the concentration in plasma.The system is specifically indicated as an aid to monitor the effectiveness of diabetes control. It is not for diagnosis of or screening for diabetes.

4.3 The selected comparison method

The selected comparison method is a fully specified method which, in the absence of a reference method, serves as the common basis for the comparison of a field method.
The selected comparison method should be an established hospital laboratory method and is used for determining the bias and accuracy of the results from the evaluated method.

The selected comparison method in this evaluation was the routine method for P—Glucose in the Department for Clinical Chemistry at Karolinska University Hospital in Huddinge. It is a Roche hexokinase method, Gluco-quant Glukos/HK, applied on a Modular Analytics P instrument from Roche Diagnostics. This method is put in practice completely according to the instructions from Roche.

The selected comparison method in this evaluation is below called "the comparison method" and is described more in details in Attachment 2.

4.3.1 Verification of the comparison method

The verification of the comparison method is described in Attachment 2. In summary the verification showed that:

- The imprecision of the comparison method calculated from the duplicate measurements on patient samples, was approximately 1 CV% except for the "very low" level. For the "very low" P—Glucose level, below 2,5 mmol/L, the CV was 4,4%. The CV for the internal quality control results was maximum 2,5%. The imprecision figures of the comparison method are considered to be good and normal for a hospital laboratory method for P—Glucose.
- The bias of the comparison method was eliminated by recalibration of the comparison method results with reference material from NIST (SRM 965b). The comparison method measurements in the evaluation were performed in two series, one short series early in the evaluation and one with 90% of the results at the end of the evaluation. The two series of measurements were recalibrated separately. The recalibrations were adjustments of the routine calibration normally performed with calibrators from Roche Diagnostics. The adjustments were carried out by means of ordinary linear regression by the following adjustment equations:
  for the short series run Adjusted value = 0,9775 × Unadjusted value − 0,0029 and
  for the long series run Adjusted value = 1,0068 × Unadjusted value − 0,0517

4.4 Planning of the evaluation

SKUP in Sweden received request for an evaluation of the StatStrip measuring system from Johan Vikner, representative of the supplier Menarini. At the time of the request, no StatStrip system had been sold on the Scandinavian market.

The protocol for the evaluation was drawn up during autumn of 2010, based on the guidelines: *"Evaluation of analytic instruments. Guidelines particularly designed for evaluation of instruments in primary health care"* [12]. The measurements in the evaluation were carried out during 2011. The evaluation is a complete evaluation according to the SKUP guidelines.

The evaluation comprised the following studies:

In a hospital laboratory:

- Compilation of facts about the measurement system
- Determination of repeatability with 100 arterial patient samples
- Comparison with an established hospital laboratory method for P—Glucose
- Evaluation of user-friendliness
- Interference studies. Effect of varying haematocrit, pH, pO2 and sodium concentrations in the samples.

In hospital wards:

- Determination of repeatability with 40 patient samples at each unit
- Comparison with an established hospital laboratory method for P—Glucose
- Evaluation of user-friendliness

After an inquiry from SKUP, the Department of Clinical Chemistry within Karolinska University Laboratory, in Stockholm, Sweden, agreed to perform the hospital laboratory part of the evaluation.

The evaluation among the intended users was carried out at three hospital wards at Karolinska University Hospital in Huddinge.

Before the evaluation, Arne Mårtensson from SKUP in Sweden drafted the preliminary protocol in collaboration with colleagues within SKUP, Nils Uhlin and Johan Vikner from Menarini and the involved persons at Karolinska University Hospital.

At the start-up meeting in January 2011, at the Department of Clinical Chemistry in Huddinge/Stockholm, the protocol was also thoroughly discussed and finally agreed upon.

Contracts were made between SKUP and the Department of Clinical Chemistry within Karolinska University Laboratory, between SKUP and the hospital wards at Karolinska University Hospital in Huddinge and between SKUP and Menarini.

Arne Mårtensson has compiled this report. The report has been sent to colleagues within SKUP in Sweden, Denmark and Norway and to Menarini and Nova Biomedical. They have all discussed and commented on the report and influenced this final report.

4.4.1 Evaluation sites and persons involved

This evaluation was carried out both in a hospital laboratory by an experienced biomedical laboratory scientist (medical technologist) under conditions when it is most likely to perform well and under real-life conditions in the hands of the intended users, the personnel at two hospital wards, with limited laboratory experience. This evaluation of StatStrip is thus a complete SKUP evaluation.

4.4.1.1 Organisation

Karolinska University Hospital comprises two large hospitals within the Stockholm area in Sweden, one in Solna and one in Huddinge. The two hospital laboratories at these two hospitals and many of the other clinical laboratories within the Stockholm area are organised within the Karolinska University Laboratory.

4.4.1.2 Participating hospital laboratory

The hospital laboratory evaluation took place at the Intensive-Care Unit in Huddinge with samples from the blood gas instrument. Among the samples already collected for blood gas analysis, samples with concentrations all over the measuring range of StatStrip could be selected. The samples were representative for samples in intensive-care.

One biomedical laboratory scientist did all the measurements with StatStrip in the hospital laboratory evaluation.

4.4.1.3 Participating hospital wards

The following hospital wards at the Karolinska University Hospital in Huddinge were planned to participate in the evaluation of StatStrip:

The Maternity Ward
The Neonatal Intensive-Care Unit
The Day-Care Unit M71 of the Endocrine Clinic

Nurses/midwifes did all the measurements with StatStrip for the evaluation at the hospital wards.

4.4.1.4 Participating persons

Table 3 contains an overview of the persons involved in the evaluation, and their respective responsibility.

Table 3. Persons responsible for various parts of this evaluation

Gunilla Forslund	Biomedical Laboratory Scientist and Laboratory Instructor	Local leader of the evaluation at the Karolinska University Hospital, site Huddinge. Carried out the measurements on StatStrip in the hospital laboratory part of the evaluation. This part was performed at Intensive-Care Unit.
Björn Garpefjord	Clinical Biochemist	Responsible for the measurements with the comparison method in the Department of Clinical Chemistry at the Karolinska University Hospital, site Huddinge.
Manijeh Gharihaghighat,	Head Nurse	Contact person for the evaluation at the Maternity Ward
Ingrid Erling Paula Kurth Mahvush Teshnizi	Midwifes	Carried out the measurements on StatStrip at the Maternity Ward.
Hilkka Lahnalampi	Diabetes Nurse	Contact person for the evaluation at the Endocrine Clinic
Maria Segerström	Nurse	Carried together with Hilkka Lahnalampi out the measurements on StatStrip at the Endocrine Clinic during the evaluation.
Nils Uhlin	Product Manager	Contact person at Menarini in Sweden before and during the evaluation.
Johan Vikner	Business Unit Manager	Partner in the discussion of the protocol for the evaluation. Representative for Menarini in Sweden.
Andrei Malic	European Director of Medical and Scientific Affairs	Partner in the discussion of the protocol for the evaluation. Representative for Nova Biomedical.
Euan Donald	POCT Product Manager	Partner in the discussion of the protocol for the evaluation. Representative for Nova Biomedical UK.
Arne Mårtensson	Clinical Biochemist	Organiser of the evaluation. Author of this report. Co-ordinator of SKUP in Sweden
Lena Morgan	Biomedical Laboratory Scientist	Assistant co-ordinator of SKUP in Sweden

4.5 Evaluation procedure

4.5.1 Procedures common for all evaluation sites

4.5.1.1 Training

Menarini in Sweden was responsible for the training in operation of StatStrip. Training was provided by Nils Uhlin for those who were going to do the hands-on work with StatStrip. The training session was similar to what is normally done when the system is sold to a new customer. The duration of the session was less than one hour. When the evaluation began, the biomedical laboratory scientist repeated the training, but then the evaluators managed the instruments single-handedly, without any supervision or correction from the biomedical laboratory scientist.

4.5.1.2 Agreement between different StatStrip meters

The parallel evaluation in the hospital laboratory and at the hospital wards required four StatStrip meters. Calibration agreement between the used instruments was checked before the evaluation. See Attachment 1

4.5.1.3 StatStrip test strips

Three different lots of test strips were used in the evaluation. Approximately one third of the measurements were performed with each lot at each evaluation site. The two measurements in each duplicate were performed with the same lot.

4.5.1.4 Samples

The samples in the hospital laboratory and at the Maternity ward were collected in tubes and mixed before measurements according to procedures in the ISO standard.

The capillary samples at the Endocrine clinic were collected according to the SKUP procedures used for evaluations of glucose meters for self-monitoring. That means that the samples collected and measured directly from multiple capillary finger punctures.

4.5.1.5 Handling of samples for the comparison method

Samples to the comparison method were collected in Microvette Li-heparin tubes (300 µL) from Sarstedt. Immediately after collection, the samples were centrifuged for three minutes at 10 000 g (alternatively in 10 minutes at 2000 g), and the plasma was separated immediately into a test tube made for freezer storage. Each sample was frozen immediately and was kept frozen until measured.

The capillary samples collected for the comparison method in the evaluation were kept frozen in a −20 °C freezer close to the collection place. At least once a week, the hospital laboratory personnel transferred the frozen samples from the −20 °C freezer to a −70 °C freezer in the laboratory.

The samples were measured with the comparison method in two series within a few months from collection and were thawed and well mixed immediately before the measurements. The first series were measured after 18 samples had been collected in the hospital laboratory. The second series of measurements were performed when all samples had been collected. Required minimum volume in the sample cup for duplicate measurement with the comparison method was 100 µL. Each sample was measured in duplicate. Both series of measurements included reference materials from NIST.

4.5.1.6 Handling of results etc.

All results were registered consecutively by the persons doing the practical work in the evaluation. All recordings were signed.

Other data was also recorded e.g. date and time of measurement, opening of a new bottle of control, serial number of each used instrument.

4.5.1.7 Quality Control

StatStrip automatically performs an electronic self-test of the meter and the test strip at each measurement, as described in Attachment 1, Table 4, Quality control.

Daily internal quality control measurements were carried out throughout the evaluation period at all evaluation sites. Control solutions for StatStrip supplied by Menarini were used. On each day of analysis controls on two levels were analysed on each instrument.

4.5.1.8 User-friendliness

The users of StatStrip in the evaluation also evaluated the user-friendliness. The user-friendliness was evaluated during and immediately after the practical work, using a questionnaire drafted by SKUP. The questionnaire was translated into Swedish and was adapted to this evaluation before being used.

4.5.2 Evaluation procedure in the hospital laboratory

4.5.2.1 Conditions

The "standardised conditions" are chosen to give the evaluated measurement system best opportunities for good performance. The equipment is handled by personnel well qualified in clinical laboratory work.

4.5.2.2 Selection of samples

100 patient samples are included. Arterial blood collected in the Intensive-Care Unit was used. The samples for blood gas analyses were collected in syringes with lithium heparin (or lithium/sodium heparin). The measurements for the evaluation were performed within two hours from the sample collection. The P—Glucose results and the haematocrit results from the blood gas analyzer were used for selection of samples for the evaluation with varying concentrations. Samples were selected to cover the StatStrip measuring range (P—Glucose 0,6 — 33,3 mmol/L) and also to include samples with low and high haematocrit results.

After 50 consecutive samples had been measured the aim was to obtain an even distribution of concentration results over the StatStrip measuring range. A minimum of five samples with P—Glucose <3,0 mmol/L and five samples with P—Glucose >20,0 mmol/L were pursued in the evaluation.

As low/normal concentrations were lacking, samples with low concentration were produced by just letting blood gas sample with normal P—Glucose concentration stand overnight in room temperature.

As high concentrations were also lacking, samples with high concentration were produced by adding small volumes of glucose solutions to blood gas samples with normal P—Glucose.

4.5.2.3 Handling of samples and measurements

Each sample was first mixed carefully, at least three minutes, before measured in duplicate on StatStrip. Both measurements on the same patient sample were performed with the same StatStrip meter. The sample was then immediately transferred to a tube without additive and centrifuged to produce plasma for the comparison method. See the procedure for the comparison method samples in section4.5.1.2.

Both measurements of a sample were performed with test strips from the same lot number. Each evaluation day the lot of test strips were changed so the three lots were used under the same period of days and in similar numbers.

The measurements in the hospital laboratory were performed on at least 20 different days.

4.5.2.4 Handling of results
See section 4.5.1.6.

4.5.2.5 Internal quality control
See section 4.5.1.7.

4.5.2.6 Precision
Repeatability for StatStrip was calculated from duplicate results of the 102 arterial samples.

4.5.2.7 Comparison of methods
The comparison of methods was carried out with the results from the 102 arterial samples measured in duplicates with the StatStrip system and with the comparison method. The mean deviation (bias) with confidence interval was calculated for all results, and for the results divided into three concentration intervals. The first single results are presented in a difference plot to evaluate the StatStrip accuracy.

4.5.2.8 User-friendliness
See section 4.5.1.8.

4.5.2.9 Supplementary evaluations
This evaluation does not include any systematic interference study. However, the samples tested in the hospital laboratory part of the evaluation derive from patients in the Intensive-Care Unit. Many of these samples contain drugs which are in common use in an Intensive-Care Unit and the found analytical quality is probably typical for what will be achieved in intensive care.

The manufacturer was asked to check that the calibration agreement between the lots used in this evaluation is typical for marketed lots. This evaluation does not include any further investigation of lot-to-lot-variation between different lots of test strips. However the evaluation was made with three lots, and the results will show the reality, where several lots are used at the same time in the market. A simple check of the agreement between the different lots used was done by visual inspection of the difference plot for the arterial samples in the hospital laboratory. See section 5.3.2

The manufacturer/requesting company was asked to check that the agreement between the StatStrip meters used in this evaluation was typical for marketed meters. This evaluation just included a check of the instrument-to-instrument-variation between the meters used in the evaluation performed before the evaluation started. See section 5.3.1

4.5.3 Evaluation in two hospital wards

4.5.3.1 Conditions
The evaluation within the hospital wards shows how StatStrip work under real conditions, when used by the intended users.

The evaluation in the hospital wards was guided by an experienced biomedical laboratory scientist. The measurements in the evaluation were performed as if they were a part of the everyday life in the hospital wards.

4.5.3.2 Recruitment of patients/sampling

In the Maternity Ward:

When newborn children in the Maternity Ward are 48 h old they are all screened for genetic disorders. For this purpose a venous sample is collected by the personnel in the Maternity Ward. When sampling the blood flows freely from an injection needle into a micro tube. The frequency of these samples was about 20 per day. At the same sample collection occasion 300 µL extra venous blood was collected and used for the evaluation. A prerequisite was however that the parents had been informed by a letter in which it was explained that participation in the evaluation was voluntary. Verbal consent was considered to be sufficient as no extra sampling occasion was needed for the evaluation. If there was problem to obtain the extra blood from a child, the sample collection on that child was discontinued. The blood was collected in 300 µL marked tube with lithium heparin as additive. The biomedical laboratory scientist was present at the ward and immediately placed one large drop of blood from the tube on a piece of Parafilm. A nurse at the Maternity Ward performed duplicate measurements with StatStrip by adding sample to the test strips from the drop.
The biomedical laboratory scientist then directly centrifuged the sample to produce plasma for the comparison method, see section 4.5.1.4. There were no requirements on the P—Glucose-concentrations in the collected samples so the samples were included in the evaluation consecutively as they were available.

In the hospital ward for Neonatal intensive-care:

The following was planned: Normally a child in the Neonatal ward, which needs a check of the P—Glucose level, is checked with a capillary sample measured with Abbott FreeStyle Lite measurement system. For the evaluation, capillary samples were planned to be collected for both FreeStyle Lite and StatStrip on 80 children. Every second child should be checked with duplicate measurements on FreeStyle Lite and every second child should be checked with duplicate measurements on StatStrip. If there was problem to obtain the extra blood from a child, the sample collection on that child should be interrupted.

The above planning could not be implemented as the nurses in the Neonatal Intensive-Care Unit thought it was too difficult to achieve the required blood volume. The evaluation at this department was then discontinued.

In the Endocrine Clinic

The P—Glucose-levels of the diabetes patients at the Day-Care Unit M71 of the Endocrine Clinic are usually checked with a single capillary sample measured with a routinely used glucose measurement system. For the evaluation extra capillary sampling was necessary on 40 patients. Participation in the evaluation was voluntary. Verbal consent was considered to be sufficient as no other measurands than the requested were measured. The first drop of blood from each skin puncture was wiped off. On each patient there were three skin punctures and four sample collections for the different methods performed in the following order:

1. Skin puncture 1.
2. Routinely used glucose measurement system.
3. StatStrip 1.
4. Skin puncture 2.
5. Comparison method
6. Skin puncture 3.
7. StatStrip 2.

The three skin punctures on one patient were performed close to each other and on the same finger. The whole sampling sequence for one patient was typically finalised within two minutes.

The sample for the routinely used glucose measurement system was collected, by a diabetes nurse in the clinic, directly into the sample collection device for routinely used glucose measurement system and was measured immediately after collection.

The capillary samples for StatStrip test were applied directly on the test strips, by a diabetes nurse in the clinic. The two samples from each patient were always collected and measured by the same person and with the same StatStrip meter.

A 300 µL capillary sample for the comparison method was collected by the biomedical laboratory scientist. See section 4.5.1.4 for description of the procedure for the comparison method samples.

The capillary sampling routine described above is close to the routine used in other SKUP evaluations of glucose meters and was outlined already in the agreed protocol for this evaluation. SKUP followed the intention in the protocol to evaluate also the analytical quality with capillary sampling as many hospital wards use capillary sampling for near-patient glucose measurements. The observed variation in measurement results partly originate from different glucose concentration in different capillary drops [13].

In all wards

In the evaluation in the wards there were no requirements on the P—Glucose-concentrations in the collected samples so the patients that participated in the evaluation were chosen consecutively or randomly as random concentrations were fit for purpose. The samples were collected over at least five different days and within 3-4 weeks.

Test-strips from three different lots were used. The duplicate measurements on the first third of samples were measured with one lot, the second third of samples were measured with a second lot and the last third of samples were measured with a third lot.

4.5.3.3 Handling of results

See section 4.5.1.6.

4.5.3.4 Quality control

See section 4.5.1.7.

4.5.3.5 Precision

Repeatability of StatStrip was calculated from the duplicate results of the 40 samples measured at each hospital ward.

4.5.3.6 Comparison of methods

The comparison of methods was carried out with the results from the 40 patients from each hospital ward measured in duplicate on StatStrip and on the comparison method.

The results are presented in a difference plot. The mean deviation (bias) with confidence interval was calculated for all results and for the results divided into two concentration level groups.

4.5.3.7 User-friendliness

See section 4.5.1.8.

4.6 Statistical terms, expressions and calculations

The statistical terms, expressions and calculations used by SKUP is described in Attachment 3. The attachment is a short extract of the comprehensive SKUP-document "Statistics in SKUP reports", available at the SKUP website [14].

4.7 Additional equipment and product details

Additional equipment and product details for used equipment are presented in Attachment 1.

5 Results and discussion

5.1 Analytical quality of StatStrip in the hospital laboratory

5.1.1 Missing and excluded results and check calculations

See Attachment 4 Table 4A.

5.1.2 Imprecision of StatStrip in the hospital laboratory

5.1.2.1 Repeatability of StatStrip with arterial patient samples

Results from arterial patient samples were sorted and divided into four level groups according to duplicate mean concentrations of the comparison method. The wide concentration interval in this evaluation and the different repeatability CV at the very low level justified a division in four level groups. The repeatability was calculated from the duplicate StatStrip results.
The outlier test was performed on relative differences (d/m) and formula 2 described in Attachment 3, section 2.2 was used for the calculation of imprecision. See Table 4.

Table 4. Repeatability of StatStrip with arterial patient samples in the hospital laboratory

Level	Comparison method interval (mmol/L)	n	Excluded results	StatStrip mean (mmol/L)	CV%[#] (95% confidence interval)
Very low	1,0 — 2,5	8	0	1,8	6,7 (4,2 — 12,8)
Low	4,7 — 6,8	19	0	6,4	2,6 (2,0 — 3,9)
Medium	6,9 — 10,4	33	0	8,1	3,1 (2,5 — 4,1)
High	10,6 — 29,2	28	0	18,1	2,7 (2,2 — 3,7)

[#] The calculated CV values are practically measures of repeatability, but they also include some additional variance components arising from changes in conditions during the collection of measurement data: three different batches of test strips were used, the concentrations in the samples were varying, the samples had varying matrix and the different samples were measured different days.

5.1.2.2 Internal quality check with control solutions

The daily internal quality control results were used to check that the StatStrip systems used in evaluation worked properly. All results from these measurements were inside the acceptance limits set by the manufacturer. The internal quality control results were used for calculation of the intermediate imprecision. See Attachment 1 section 4.1. However, it must be remembered that the results with control solutions not always reflect the results with patient samples.

5.1.2.3 Assessment of the imprecision of StatStrip

The two samples for StatStrip in this part of the evaluation were collected from the same tube of arterial blood. The sampling errors were thereby minimized.

According to quality goal set up by SKUP, the repeatability CV of StatStrip should not exceed 4%. The CV-values were around 3% for all results except for the results in the "very low" level group with values below 2,5 mmol/L. No differences between the duplicates in the "very low" group exceeded 0,2 mmol/L, which is considered as good. After the evaluation it was realised that the quality goal for imprecision was unrealistic for this level group. Not even the comparison method fulfilled the quality goal for imprecision at this concentration level.

The precision of StatStrip with arterial blood in the hospital laboratory was assessed as good and fulfilled the quality goal.

5.1.3 Bias of StatStrip in the hospital laboratory

5.1.3.1 Bias of StatStrip with arterial patient samples

Results from 86 arterial patient samples were sorted and divided into three level groups according to the duplicate mean concentrations of the comparison method. The bias in each level group was calculated from the means of the duplicate sample results of StatStrip compared with the means of the duplicate determinations with the comparison method. See Table 5.

Table 5. StatStrip bias with arterial patient samples

Level group	Comparison method interval (mmol/L)	n	Number of excluded results	StatStrip method mean (mmol/L)	Bias (95% confidence interval) (mmol/L)	Bias (95% confidence interval) (%)
Low	1,0 — 6,8	27	0	5,0	+0,28 (+0,15 — +0,41)	+6,8 (+3,9 — +9,6)
Medium	6,9 — 10,4	32	0	8,2	+0,14 (+0,01 — +0,26)	+1,7 (+0,2 — +3,3)
High	10,6 — 29,2	27	0	18,2	+0,20 (−0,05 — +0,46)	+1,1 (−0,3 — +2,5)
All	1,0 — 29,2	86	0	10,3	+0,20 (+0,10 — +0,30)	Not calculated

5.1.3.2 Assessment of the bias with arterial patient samples

Expressed in mmol/L StatStrip showed similar bias over the whole measuring range: +0,2 mmol/L. This bias is assessed as acceptable.

5.1.4 Accuracy of StatStrip in the hospital laboratory

5.1.4.1 Accuracy of StatStrip with arterial patient samples

There were 97 arterial sample results for the estimation of accuracy. Seven samples showed the result code "LO <0,6" interpreted as <0,6 mmol/L and two samples showed the result code "HI >33,3", interpreted as >33,3 mmol/L. The samples with the StatStrip result code "LO <0,6" got the following results with the comparison method <0,10, <0,10, 0,13, 0,29, 0,44, 0,48, 0,58 and 0,84 mmol/L. The samples with the StatStrip result code "HI >33,3" got the following results with the comparison method 33,7 and 34,9 mmol/L. There was also one sample with the StatStrip result code "HI >33,3" that had already been excluded because of too large difference between the comparison method duplicate results 55,0 and 52,6 mmol/L. All the result codes were confirmed as correct and these results are counted as successful measurements in the evaluation but they are not shown in the plot in figure 2.

The agreement between StatStrip results and comparison method results with arterial samples is illustrated in a difference plot, Figure 2. In the plot the x-axis represents the mean result of the duplicate measurements with the comparison method. The y-axis shows the deviation of the first measurement on StatStrip from the mean value of the duplicate results of the comparison method. The difference plot illustrates both random and systematic deviations and reflects the accuracy of StatStrip.

All results except one fell inside the quality goal limits. One result fell just above the high quality goal limit.

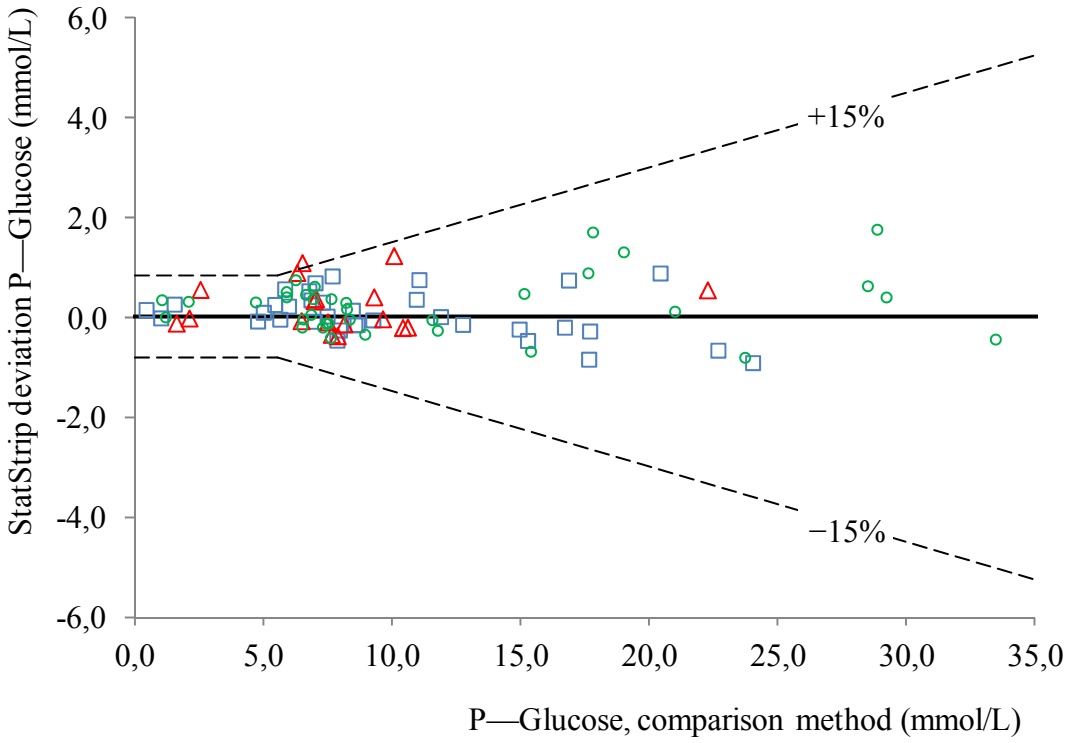


Figure 2. Difference plot, arterial patient samples in the hospital laboratory
The deviation of the arterial StatStrip result from the arterial comparison method result is shown for 88 patients.
Results are shown with different symbols depending on used lot of test strips:
□ blue squares lot 310144249 △ red triangles lot 310193249 ○ green circles lot 310223249.

5.1.4.2 Assessment of the accuracy of StatStrip with arterial patient samples

In the hospital laboratory 96 out of 97 results or 99%, were inside the quality goals limits. With arterial samples in the hospital laboratory the StatStrip results thus fulfilled the quality goal for accuracy.

5.1.5 Influence of some possible interferents

5.1.5.1 Influence of B—Haematocrit

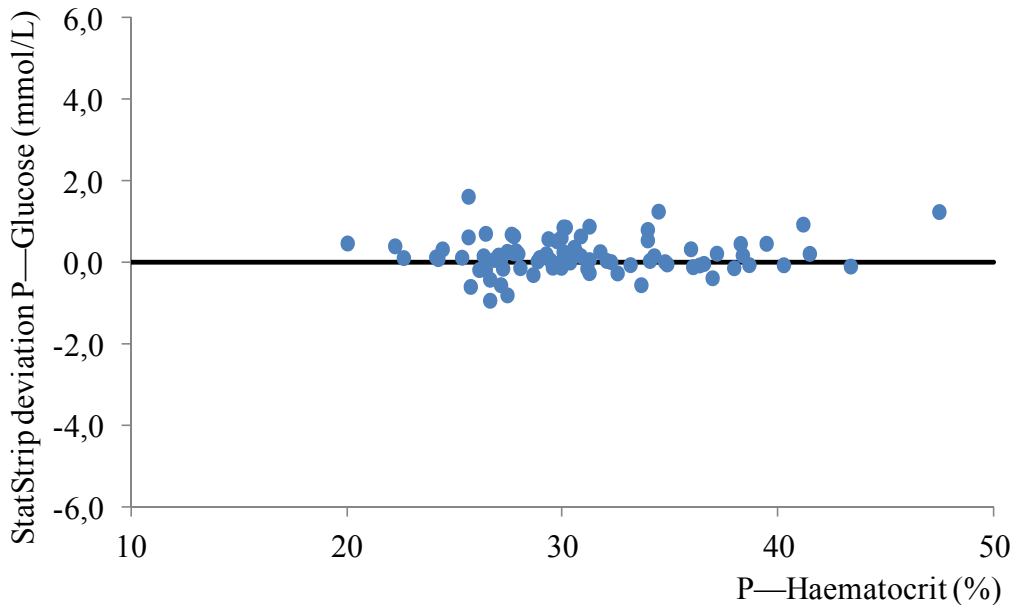


Figure 3. StatStrip results with arterial samples. Influence of B—Haematocrit
The deviations of StatStrip results in relation to the B—Haematocrit of the samples are shown for 82 arterial patient sample results.

5.1.5.2 Assessment of the influence of B—Haematocrit on the StatStrip results

As can be seen in Figure 3 the trend line is described by the equation y = 0,0087x - 0,0902 The slope of the trend line, +0,009 with the 95% confidence interval −0,010 to +0,028, is not significantly different from zero. The conclusion is that the StatStrip results were not influenced by haematocrit in the samples within the examined haematocrit interval 20 to 47%.

In Sweden the haematocrit influence is sometimes expressed as change in P—Glucose result per 10% change in haematocrit. For StatStrip this value with the 95% confidence interval is +0,09 (−0,10 to +0,28).

5.1.5.3 Influence of B—pH

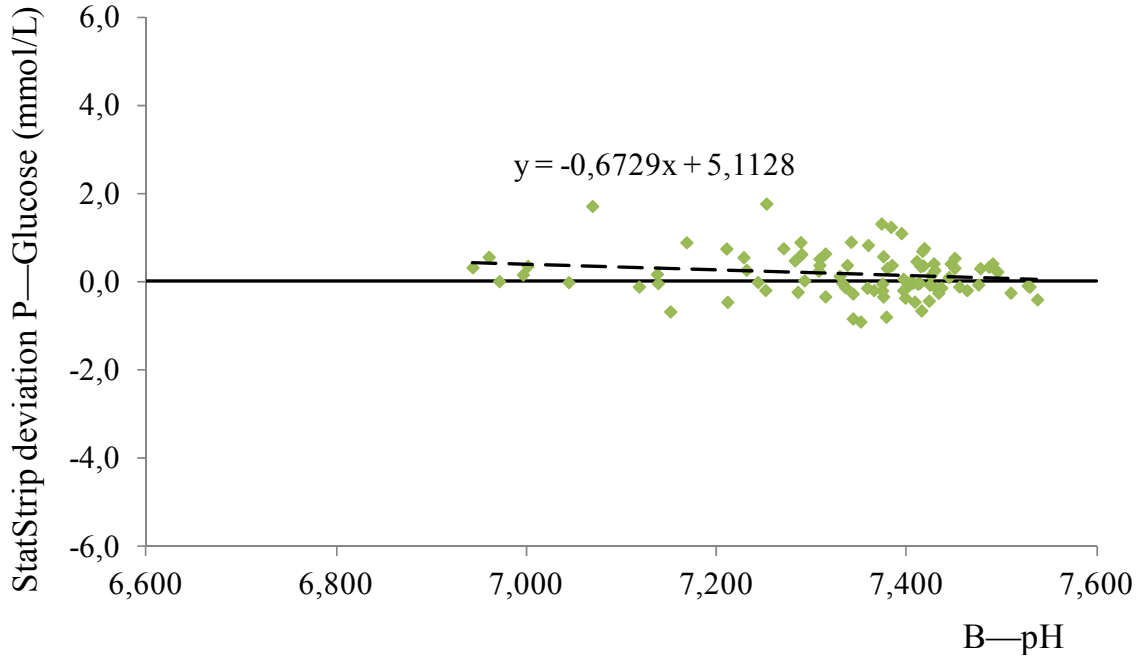


Figure 4. StatStrip results with arterial samples. Influence of B—pH
The deviations of StatStrip results in relation to the B—pH activity in the samples are shown for 89 arterial patient sample results.

5.1.5.4 Assessment of the influence of B—pH on the StatStrip results

As can be seen in Figure 4 the trend line is described by the equation y = −0,6729x + 5,1128 The slope of the trend line, −0,67 with the 95% confidence interval −1,46 to +0,12, is not significantly different from zero. The conclusion is that the StatStrip results were not influenced by the B—pH activity in the samples within the examined B—pH interval 6,9 to 7,5.

5.1.5.5 Influence of B—pO_2

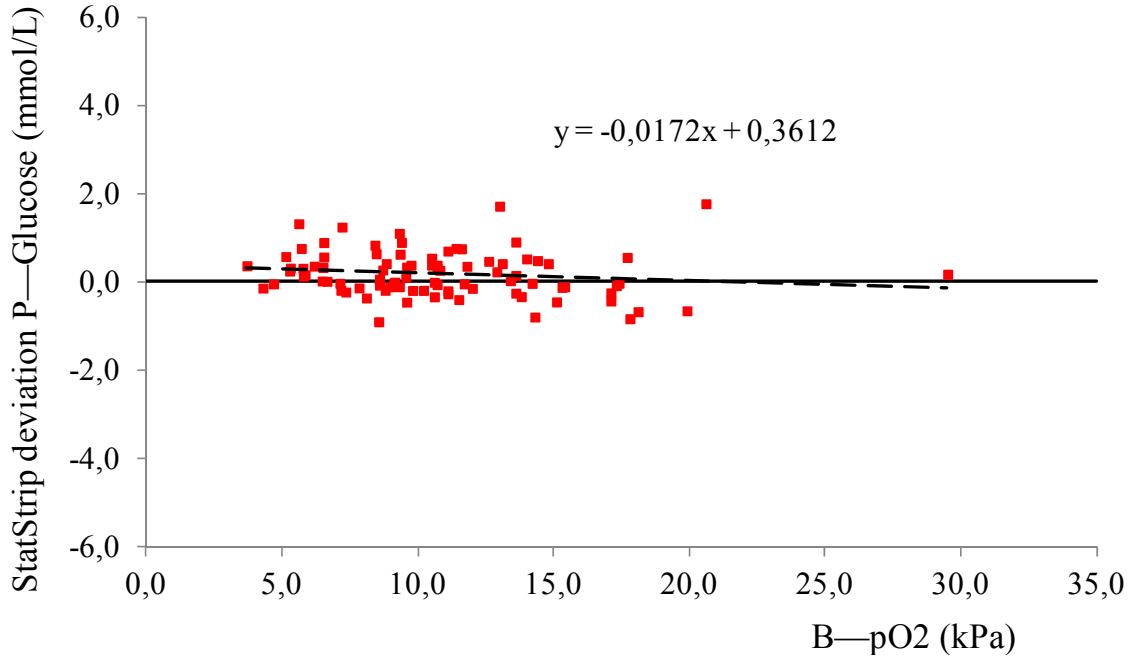


Figure 5. StatStrip results with arterial samples. Influence of B—pO_2
The deviations of StatStrip results in relation to the B—pO_2 partial pressure in the samples are shown for 89 arterial patient sample results.

5.1.5.6 Assessment of the influence of B—pO_2 on the StatStrip results

As can be seen in Figure 5 the trend line is described by the equation y = −0,0172x + 0,3612 The slope of the trend line, −0,017 with the 95% confidence interval −0,042 to +0,008, is not significantly different from zero. The conclusion is that the StatStrip results were not influenced by the B—pO_2 partial pressure in the samples within the examined pO_2 interval 4 to 30 kPa.

5.1.5.7 Influence of P—Sodium

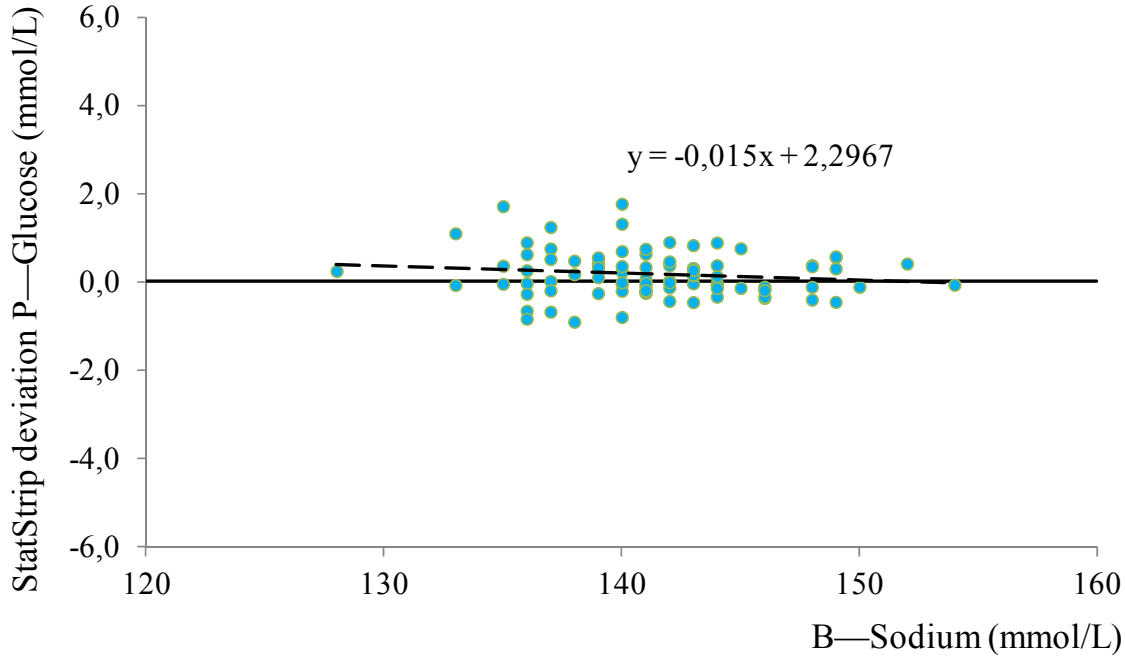


Figure 6. StatStrip results with arterial samples. Influence of P—Sodium
The deviations of StatStrip results in relation to the P—Sodium concentration in the samples are shown for 89 arterial patient sample results.

5.1.5.8 Assessment of the influence of P—Sodium on the StatStrip results

As can be seen in Figure 6 the trend line is described by the equation y = −0,015x + 2,2967 The slope of the trend line, −0,015 with the 95% confidence interval −0,039 to +0,009, is not significantly different from zero. The conclusion is that the StatStrip results were not influenced by the P—Sodium concentration in the samples within the examined P—Sodium interval 128 to 154 mmol/L.

5.2 Analytical quality of StatStrip at the hospital wards

5.2.1 Missing and excluded results and check calculations

See Attachment 4 Table 6A and Table 6B.

5.2.2 Imprecision of StatStrip at the hospital wards

5.2.2.1 Repeatability of StatStrip with patient samples

For each of the two hospital wards the results from about 40 patient samples were sorted and divided into two level groups according to the duplicate mean concentrations of the comparison method. The outlier test was performed on relative differences (d/m) and formula 2 described in Attachment 3, section 2.2 was used for the calculation. The repeatability within each level group was calculated from the duplicate StatStrip results from each hospital ward. See Table 6 and 7.

Table 6. Repeatability of StatStrip with venous patient samples in the Maternity Ward

Level	Comparison method interval (mmol/L)	n	Number of excluded results	StatStrip mean (mmol/L)	CV* (95% confidence interval) (%)
Low	3,1 — 4,5	21	0	3,9	7,2 (5,5 — 10,4)
Medium	4,5 — 6,4	20	0	4,6	5,0 (3,8 — 7,3)

* The calculated CV values are practically measures of repeatability, but they also include some additional variance components arising from changes in conditions during the collection of measurement data: three different lots of test strips were used, varying concentrations and varying matrix in the samples and varying days between the measurements.

Table 7. Repeatability of StatStrip with capillary patient samples in the Endocrine Clinic

Level	Comparison method interval (mmol/L)	n	Number of excluded results	StatStrip mean (mmol/L)	CV* (95% confidence interval) (%)
Medium	4,9 — 10,0	20	0	7,5	7,1 (5,4 — 10,3)
High	10,0 — 27,2	19	0	15,9	—

* See below table 6.

There is no CV calculated for Endocrine Clinic results in the high concentration level. A prerequisite for the formula used is that there is no statistically significant difference between the sum of the first and the sum of the second results in the duplicates. The check calculations in Attachment 4 Table 4b show however that the first value minus the second value in each duplicate result produced a mean difference of +0,73 mmol/L and the confidence interval for the mean difference did not include zero difference. These differences seem to be influenced not only by random errors but also a systematic difference that makes the data unsuitable for calculation of a "true imprecision CV". This significant difference was found only in this result group.

5.2.2.2 Assessment of the imprecision of StatStrip with patient samples

The two samples for StatStrip in the Maternity ward evaluation were collected from the same tube of venous blood. The sampling errors were thereby minimized. On the other hand, the two samples for StatStrip in the Endocrine clinic evaluation were collected directly from the finger punctures. The sampling errors were thereby included in the results and probably explain that the repeatability figures are higher in the Endocrine clinic.

For the patient samples measured at the two hospital wards, the repeatability CV was between 5,0% and 7,2%. These figures are similar to the imprecision specifications given by Nova Biomedical. See Attachment 1, section 1.2.

In the hospital wards StatStrip did not fulfil the SKUP quality goal for imprecision. At the Maternity Ward, the results within the concentration interval 4,5 to 6,4 mmol/L were inconclusive on fulfilling the quality goal but most likely the quality goal was not fulfilled.

5.2.3 Bias of StatStrip at the hospital wards

Bias was calculated from the measurement results with samples from about 40 patients at each hospital ward. The results from each ward were sorted and divided into two level groups according to the duplicate mean concentrations of the comparison method. The bias was calculated from the means of the duplicate sample results of StatStrip compared with the means of the duplicate determinations with the comparison method in the level groups. The bias values for all results together were not calculated as there were considerable differences in bias between the level groups. See Table 8.

Table 8. Bias of StatStrip with patient samples

Level	Comparison method interval (mmol/L)	n*	Number of excluded results	StatStrip mean (mmol/L)	Bias (95% confidence interval) (mmol/L)	Bias (95% confidence interval) (%)
Maternity Ward, venous samples:						
Low	2,5 — 4,2	20	0	4,1	+0,32 (+0,15 — +0,49)	+8,7 (+4,1 — +13,4)
Medium	4,2 — 5,5	21	1	4,8	+0,05 (−0,09 — +0,19)	+1,2 (−1,9 — +4,2)
Endocrine Clinic, capillary samples:						
Medium	4,9 — 10,0	21	0	7,6	+0,34 (+0,23 — +0,44)	+4,9 (+3,4 — +6,5)
High	10,0 — 27,2	19	0	15,9	+0,54 (+0,14 — +0,94)	+3,5 (+1,1 — +6,0)

* The given numbers of results (n) are counted before exclusion of outliers. Mean and bias are calculated after exclusion of outliers.

5.2.3.1 Assessment of the bias with patient samples

StatStrip showed positive bias with the patient samples in the hospital wards. The positive bias in the hospital wards was on average +0,3 mmol/L. The achieved bias is assessed as acceptable.

5.2.4 Accuracy of StatStrip at the hospital wards

The agreement between results measured with StatStrip at the hospital wards and results from the comparison method is illustrated in a difference plot, Figure 7. In the plot the x-axis represents the mean value of the duplicate results from the comparison method. The y-axis shows the deviation of the first measurement of StatStrip from the mean value of the duplicate results of the comparison method. The difference plot illustrates both random and systematic deviations and reflects the accuracy of StatStrip.

The results in the Maternity Ward derive from healthy newborn children approximately 48 hours old. The samples can be assumed to have haematocrit values within the reference interval for infants of this age. According to Jopling et al [15] the reference interval for haematocrit in 48 hours old infants, after 35–42 weeks' gestation, is 38 to 63%. Such high haematocrit values were a challenge for StatStrip in this evaluation.

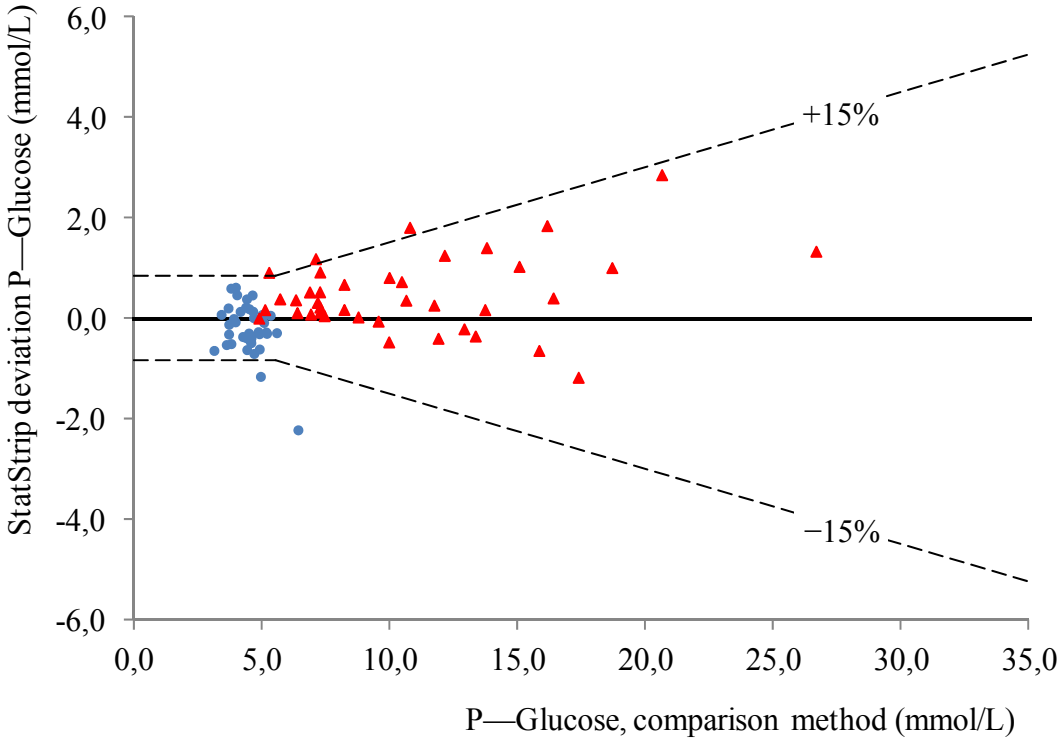


Figure 7. Difference plot, patient samples at the hospital wards
Deviation of the StatStrip result from the comparison method result are shown for 81 patient samples.
Stippled lines represent the tolerance limits ±0,83 mmol/L / ±15%.
The symbols show which hospital ward the results derive from:
- blue filled circles Newborn children at Maternity Ward, venous samples
- red filled triangles Adult persons with diabetes at Endocrine Clinic, capillary samples:

5.2.4.1 Assessment of the accuracy of StatStrip with patient samples

In the Maternity Ward 38 out of 40 venous results or 95%, were inside the quality goal limits. One of the StatStrip results outside the quality goal limits showed a large deviation of −2,2 mmol/L (−34%) from the comparison method result.

In the Endocrine Clinic 38 out of 41 capillary results or 93%, were inside the quality goal limits. The three results outside the quality goal limits were close to the limits.

The results in the hospital wards were close to the quality goal limits for accuracy and the quality goal was fulfilled in one ward and not fulfilled in the other ward.

5.3 Results valid for all evaluation sites

5.3.1 Agreement between results of different StatStrip meters

The parallel evaluation in the hospital laboratory and at the hospital wards required four StatStrip meters. Calibration agreement between the used meters was checked before the evaluation. The results and calculations of the agreement check are presented in Attachment 1, Section 3.

As can be seen in Attachment 1, Table 8, the means and CVs of all meters agreed well. There were negligible calibration differences between the StatStrip meters used at the different evaluation sites. The within-instrument CV component was 3 – 4 % and the between-instrument CV component was less than 1 %.

The requirements defined by SKUP for agreement between instruments were fulfilled. Attachment 1, Table 8 also shows which meter that was used at which site.

5.3.1.1 Assessment of the agreement between results of different StatStrip meters

The conclusion was that all meters used in the evaluation showed good calibration agreement and that variation of the results increased very little by using several meters.

5.3.2 Agreement between results of different lots of test strips

Three different lots of test strips were used in the evaluation. About one third of the measurements were performed with each lot at each evaluation site. Both duplicate measurements on each sample were performed with the same lot.

The number of tests in this SKUP evaluation was not sufficient to perform a statistical comparison between the different lots, as each patient sample was measured just with one lot.

The agreement between the different lots used was done by visual inspection of the difference plot for the arterial samples in the hospital laboratory. See section 5.1.4.1 Figure 2.

5.3.2.1 Assessment of the agreement between different lots of test strips

The evaluation showed no calibration differences between the used lots of test strips.

5.3.3 Quality goal fulfilment at all evaluation sites

The analytical quality goals used in this evaluation are specified in section 3.4. The fulfilment of the quality goals at all evaluation sites is shown in Table 9.

The SKUP quality goal for imprecision was fulfilled in the hospital laboratory but not in the hospital wards.

The SKUP quality goal for accuracy was fulfilled in the hospital laboratory measuring arterial blood from intensive care patients and in one of the two hospital wards – the one measuring venous blood from newborns.

Compared to the hospital laboratory results, the results in the hospital wards were more deviating from the comparison method results. This is due to both higher imprecision and higher positive bias.

The conditions were different in the different parts of the evaluation. In the hospital laboratory the operator was an experienced biomedical laboratory scientist and in the hospital wards the operators were midwifes or nurses.

In the hospital laboratory part of the evaluation, the samples, both for StatStrip and the comparison method were collected from the same tube of arterial blood. The sampling errors were thereby minimized.

Also in the Maternity ward part of the evaluation, the sampling errors were minimized, because the samples were collected from the same tube – in this case of venous blood.

For the Endocrine clinic part of the evaluation, capillary samples, both for StatStrip and the comparison method, were collected from three finger punctures on the same finger. The sampling errors were thereby included in the results and may explain why the repeatability and inaccuracy figures are less good for the Endocrine clinic results.

As good analytical quality is possible to achieve in the hospital laboratory the actual measurement in the meter have high quality but especially the capillary sample collection and/or sample application seems to cause more deviating results when StatStrip is used in the hospital wards.

The result of the comparison with the 10% quality goal is given for information only. The purpose with this quality goal is to graduate the performance of the real good meters. In the SKUP evaluations of meters for glucose self monitoring there are usually no results below 4,0 mmol/L so the ±0,4 mmol/L limits are not applied. The StatStrip results compared to the 10% goal are described more in detail below:

In the hospital laboratory there is 1 of 16 results outside the ±0,4 mmol/L limits and 5 of 81 results outside the ±10% limits. All results outside the limits deviate with too high StatStrip results.

At the Maternity ward there are 5 of 11 results outside the ±0,4 mmol/L limits and 8 of 30 results outside the 10% limits. 3 results are below the −0,4 mmol/L limit and 2 results are above the +0,4 mmol/L limit. 6 results are below the −10% limit and 2 results are above the +10% limit.

At the Endocrine clinic there are no results to apply the ±0,4 mmol/L limits and 9 of 40 results are outside the 10% limits. All results outside the limits deviate with too high StatStrip results.

Table 9. StatStrip results at the three evaluation sites and comparisons with different quality goals

	Repeatability	Bias		Accuracy	
Quality goal:	**≤4 CV%**	**None**	**95% within ±0,83 mmol/L / ±20% Break point: 4,2 mmol/L**	**95% within ±0,83 mmol/L / ±15% Break point: 5,6 mmol/L**	**95% within ±0,40 mmol/L / ±10% Break point: 4,0 mmol/L**
Users Category of measured patients Evaluation site Kind of sample	**Repeatability CV%**	**Bias (mmol/L)**	**Results within the above limits (%)**	**Results within the above limits (%)**	**Results within the above limits (%)**
Experienced biomedical laboratory scientist Adult intensive-care patients Intensive-Care Unit, Arterial samples	2,6 to 3,1	+0,14 to +0,28	100	99	94
Midwifes, Healthy newborn children Maternity Ward Venous samples	(5,0) to 7,2	+0,05 to +0,32	95	95	68
Nurses, Adult persons with diabetes Endocrine Clinic Capillary samples	7,1/(—)*	+0,34 to +0,54	100	93	78

The turquoise boxes show the results compared to the quality goals agreed in the protocol for this evaluation.
A result inconclusive in fulfilling the quality goal is printed within parentheses.
* (—) means an inconclusive result as the data was not suitable for calculation of the imprecision as there was a systematic difference between the duplicate results.
The second column shows the fulfilment of the agreed quality goal for imprecision.
The third column shows, for information only, the bias results.
The fourth column shows, for information only, the results assessed with the same limits for accuracy as in ISO 15197:2003.
The fifth column shows the fulfilment of the agreed quality goal for accuracy.
The sixth column shows, for information only, the results according to a 10% limits for accuracy.

5.4 Evaluation of the user-friendliness of StatStrip

At the end of the evaluation period, the users filled in a questionnaire about the user-friendliness of the StatStrip. The questionnaire and expressed opinions are presented in Table 10 to 13.

The first column explains the evaluated properties.

The second column shows the expressed rating by the users: The first row in the column shows the rating of the biomedical laboratory scientist in the hospital laboratory and the second row shows the ratings of the four nurses/midwifes at the hospital wards. In cases where the evaluated property is answered by a fact, that fact is evaluated by SKUP and the second column is left empty without ratings.

The third to fifth column show the rating options. The cells with the overall ratings from the evaluating sites are marked by thicker frames and bold text.

On the last row in each table SKUP summarises the ratings in the table.

The total rating of each row and the total rating of each table are an overall assessment of the property described on the row or in the headline of the table. A single bad rating can justify an overall bad rating if that property seriously influences on the user-friendliness of the system.

Unsatisfactory and intermediate ratings will be marked and explained below the table.

Table 10. Assessment of the information in the manual / insert

Information in manual / insert about:	Ratings	Overall rating		
		Red	**Yellow**	**Green**
General impression of the manual / insert	G Y – G G	Un-satisfactory	Intermediate	**Satisfactory**
Table of content	G G – G G	Un-satisfactory	Intermediate	**Satisfactory**
Preparations / pre-analytical procedures	G – G G R	Un-satisfactory	Intermediate	**Satisfactory**
Specimen collection	G Y G G Y	Un-satisfactory	Intermediate	**Satisfactory**
Measurement / reading	G Y G G G	Un-satisfactory	Intermediate	**Satisfactory**
Measurement principle	G Y G G Y	Un-satisfactory	Intermediate	**Satisfactory**
Sources of error	G G Y G G	Un-satisfactory	Intermediate	**Satisfactory**
Fault-tracing / troubleshooting	G G Y G G	Un-satisfactory	Intermediate	**Satisfactory**
Index*		Un-satisfactory	Intermediate	Satisfactory
Readability / clarity of presentation	G Y – G G	Un-satisfactory	Intermediate	**Satisfactory**
Available in Danish, Norwegian, Swedish		No	**Danish only on demand**	Yes
Others comments about information in the manual / insert (please specify)[#]	– – – Y R[#]	Un-satisfactory	Intermediate	Satisfactory
Rating for information in manual / insert		Un-satisfactory	Intermediate	**Satisfactory**

***** There is no index in the manual
[#] Specified in the negative comments below.

The evaluators made the following additional comments concerning the information in the manual / insert:

Positive comment:

- Good illustrations!

Negative comments:

- Too much to read (3 evaluators)

Table 11. Assessment of the time factors

Time factors	**Ratings**	**Overall rating**		
		Red	**Yellow**	**Green**
Time for preparations / pre-analytical time		>10 min	6 — 10 min	**<5 min**
Analytical time		>10 min	6 — 10 min	**<5 min**
Required training time		>8 hours	2 to 8 hours	**<2 hours**
Stability of test strips, unopened package		<3 months	3 to 5 months	>5 months
Stability of test strips, opened package		<14 days	14 to30 days	>1 month
Others comments about time factors (please specify)		Un-satisfactory	Intermediate	Satisfactory
Rating for time factors				**Satisfactory**

Table 12. Assessment of the quality control possibilities

Quality Control, possibilities to perform:	**Ratings**	**Overall rating**		
		Red	**Yellow**	**Green**
Internal quality control	G Y – G –	Un-satisfactory	Intermediate	**Satisfactory**
External quality control	– – – G –	Un-satisfactory	Intermediate	**Satisfactory**
Stability of the quality control materials, unopened		<3 months	3 to 5 months	**>5 months**
Stability of the quality control materials, opened		<1 day	<1 week	**>1 week**
Storage conditions for quality control materials, unopened		–20 °C	+2 to +8 °C	**+15 to +30 °C**
Storage conditions for quality control materials, opened		–20 °C	+2 to +8 °C	**+15 to +30 °C**
Usefulness of the quality control	G G – G –	Un-satisfactory	Intermediate	**Satisfactory**
Others comments about quality control (please specify).	– – – – –	Un-satisfactory	Intermediate	Satisfactory
Rating for quality control				**Satisfactory**

Table 13. Assessment of the operation facilities

Operation facilities	**Ratings**	**Overall rating**		
		Red	**Yellow**	**Green**
Content of the test kit. Complete?	G Y – G –	Un-satisfactory	Intermediate	**Satisfactory**
Preparations / pre-analytical procedures	G Y G G –	Un-satisfactory	Intermediate	**Satisfactory**
Application of specimen	G G G G –	Un-satisfactory	Intermediate	**Satisfactory**
Specimen volume	G G G G –	Un-satisfactory	Intermediate	**Satisfactory**
Number of procedure steps	R Y G Y –	Un-satisfactory	**Intermediate***	Satisfactory
Instrument / test strips	G Y G G –	Un-satisfactory	Intermediate	**Satisfactory**
Reading	G Y G G –	Un-satisfactory	Intermediate	**Satisfactory**
Sources of error	G – – G –	Un-satisfactory	Intermediate	**Satisfactory**
Cleaning/maintenance	G – – G –	Un-satisfactory	Intermediate	**Satisfactory**
Hygiene, when using the test	G – – G –	Un-satisfactory	Intermediate	**Satisfactory**
Storage conditions for test strips, unopened package		–20 °C	+2 to +8 °C	**+15 to +30°C**
Storage conditions for test strips, opened package		–20 °C	+2 to +8 °C	**+15 to +30°C**
Environmental aspects: waste handling		Special precautions	**Sorted waste#**	No precautions
Intended users	G G G G –	Biomedical laboratory scientist	Laboratory experience	**No laboratory experience**
Size and weight of packages	G G G G –	Un-satisfactory	Less satisfactory	**Satisfactory**
Others comments about operation:^	– R G G –	Un-satisfactory	Less satisfactory	Satisfactory
Rating for operation facility				**Satisfactory**

***** Several evaluators think that the operation of StatStrip requires too many procedure steps. See comments below.

#Environmental aspects: waste handling. The used test strips contain blood which may be infectious. This property is not different from other equipment for measuring blood glucose.

^ Specified in the comments below.

The evaluators made the following additional comments concerning the operation facilities.

Positive comments:

- Good instrument features: The operator ID and the patient ID is recorded in StatStrip. Works like a mini computer
- Small sample volume
- Simple operation

Negative comments:

- Too many procedure steps such as scanning the test strip lot number, internal quality control lot number, operator ID and patient ID *
- Too many keystrokes for operating the meter
- The battery discharges quickly. After the low battery alarm has signalled it is impossible to finish the ongoing measurement
- Difficult to find stored results
- Impossible to delete patient id number in the instrument

* Comments from SKUP: The minimum of information, that the operator has to scan or key in for each measurement, is the operator ID, patient ID or sample ID and the test strip lot number (can be static and be confirmed by pressing OK).
These mandatory procedure steps can be considered as a disadvantage as they complicate the routine measurements or as an advantage as they improve the post-analytical traceability compared to many alternative glucose meters.

5.4.1 Assessment of the user-friendliness

For most of the statements in the questionnaire StatStrip was assessed as “Satisfactory”.
In general terms, the evaluators was satisfied with the system and thought it was easy to handle. Several evaluators however think that the operation of StatStrip requires too many procedure steps.

6 About SKUP

6.1 The organisation of SKUP

Scandinavian evaluation of laboratory equipment for primary health care, *SKUP*, is a co-operative commitment of NOKLUS[1] in Norway, DAK-E[2] in Denmark, and EQUALIS[3] in Sweden. SKUP was established in 1997 at the initiative of laboratory medicine professionals in the three countries. SKUP is led by a Scandinavian *steering committee* and the secretariat is located at NOKLUS in Bergen, Norway.

The purpose of SKUP is to improve the quality of near patient testing in Scandinavia by providing objective and supplier-independent information on analytical quality and user-friendliness of laboratory equipment. This information is generated by organising SKUP *evaluations*.

SKUP offers manufacturers and suppliers evaluations of equipment for primary health care and also of devices for self-monitoring. Provided the equipment is not launched onto the Scandinavian market, it is possible to have a confidential pre-marketing evaluation. The company requesting the evaluation pays the actual testing costs and receives in return an impartial evaluation.

There are *general guidelines* for all SKUP evaluations and for each evaluation a specific *SKUP protocol* is worked out in co-operation with the manufacturer or their representatives. SKUP signs *contracts* with the requesting company and the evaluating laboratories. A *complete evaluation* requires one part performed by experienced laboratory personnel as well as one part performed by the intended users.

Each evaluation is presented in a *SKUP report* to which a unique *report code* is assigned. The code is composed of the acronym SKUP, the year and a serial number. If suppliers use the SKUP name in marketing, they have to refer to www.skup.nu and to the report code in question. For this purpose the company can use a logotype available from SKUP containing the report code.

SKUP reports are published at www.skup.nu

---

[1] NOKLUS (Norwegian Quality Improvement of Primary Care Laboratories) is an organisation founded by Kvalitetsforbedringsfond III (Quality Improvement Fund III), which is established by The Norwegian Medical Association and the Norwegian Government. NOKLUS is professionally linked to "Seksjon for Allmennmedisin" (Section for General Practice) at the University of Bergen, Norway.

[2] SKUP in Denmark is placed in Hillerød Hospital. SKUP in Denmark reports to DAK-E (Danish Quality Unit of General Practice), an organisation that is supported by KIF (Foundation for Quality and Informatics) and Faglig udvalg (Professional Committee), which both are supported by DR (The Danish Regions) and PLO (The Organisation of General Practitioners in Denmark).

[3] EQUALIS AB (External quality assurance in laboratory medicine in Sweden) is a limited company in Uppsala, Sweden, owned by "Sveriges Kommuner och Landsting" (Swedish Association of Local Authorities and Regions), "Svenska Läkaresällskapet" (Swedish Society of Medicine) and IBL (Swedish Institute of Biomedical Laboratory Science).

6.2 Contact addresses to SKUP

SKUP secretariat
Grete Monsen
+47 55 97 95 02
grete.monsen@noklus.no

SKUP in Denmark
Esther Jensen
Hillerød Hospital
Klinisk Biokemisk Afdeling
Dyrehavevej 29, indgang 16A
DK-3400 Hillerød
+45 48 29 41 76
esj@hih.regionh.dk

SKUP in Norway
Grete Monsen
Camilla Eide Jacobsen
Marianne Risa
Sverre Sandberg
NOKLUS
Boks 6165
NO-5892 Bergen
+47 55 97 95 02
grete.monsen@noklus.no
camilla.jacobsen@noklus.no
marianne.risa@noklus.no
sverre.sandberg@isf.uib.no

SKUP in Sweden
Arne Mårtensson
Lena Morgan
Gunnar Nordin
EQUALIS
Box 977
SE-751 09 Uppsala
+46 18 69 31 64
arne.martensson@equalis.se
lena.morgan@equalis.se
gunnar.nordin@equalis.se

www.SKUP.nu

7 References

1. Stöckl D, Baadenhuijsen H, Fraser CG, Libeer JC, Petersen PH, Ricos C, ”Desirable Routine Analytical Goals for Quantities Assayed in serum”. Eur J Clin Biochem 1995; 33 (3): 157 – 169.
2. American Diabetes Association. Self-monitoring of blood glucose. Diabetes Care 1996; 19 (suppl 1): 62 – 66.
3. Skeie S, Thue G, Sandberg S, “Patient-derived Quality Specifications for Instruments Used in Self-Monitoring of Blood Glucose”. Clin Chem 2001; 47 (1): 67 – 73.
4. ISO 15197 (2003). In vitro diagnostic test systems - Requirements for blood-glucose monitoring systems for self- testing in managing diabetes mellitus, ed.
5. Finfer S, Chittock DR, Su SY, Blair D, Foster D, et al. NICE-SUGAR Study Investigators, Intensive versus conventional glucose control in critically ill patients. N Engl J Med 2009; 360: 1283 – 97.
6. Mahoney JJ et al. and Cembrowski GS et al, Letters to the Editor in Clin Chem 2010, 56:10, 1643–1644
7. Cembrowski GS. Comments at FDA public meeting: clinical accuracy requirements for point of care blood glucose meters. March 17, 2010. p 56. http://www.fda.gov/downloads/MedicalDevices/NewsEvents/WorkshopsConferences/UCM208601.pdf (Accessed Nov 2010).
8. Boyd JC and Bruns DE, Quality Specifications for Glucose Meters: Assessment by Simulation Modeling of Errors in Insulin Dose, Clin Chem 2001, 47:2, 209–214.
9. Alfhei K, “Vellykket landskonferanse i NOKLUS”. Tidsskrift for den Norske Legeforening 2008; 128: p. 2636
10. National consensus document from 2007 (in Swedish). Available at www.equalis.se / /Vår verksamhet / Projekt / Glukos, nationella kvalitetsmål / /Så här noggranna ska P-Glukos-resultaten vara (pdf)
11. www.sst.dk/English/NPULaboratoryTerminology.aspx
12. Christensen N G, Monsen G, Sandberg S. Utprøving av analyseinstrumenter. En veiledning spesielt beregnet for utprøving av instrumenter for primærhelsetjenesten. Alma Mater Forlag 1997, ISBN 82-419-0230-1.
13. Kimberly MM, Caudill SP, Vesper HW, Ethridge SF, Archibold E, Porter KH, and Myers GL, Within-Person, Among-Finger Variability of Capillary Blood Glucose Measurements, Clin Chem 2004, 50:12, 2389–2391.
14. www.skup.nu / ”The SKUP evaluation” (in the left menu) / “Statistics and calculations”/”Statistics in the SKUP reports, version 1.0”.
15. Jopling J, Henry E, Wiedmeier SE and Christensen RD. Reference Ranges for Hematocrit and Blood Hemoglobin Concentration During the Neonatal Period: Data From a Multihospital Health Care System. Pediatrics, February 2009; 123:2 e333-e337; doi:10.1542/peds.2008-2654

Attachment 1.

Nova StatStrip Glucose and ß-Ketone Hospital Meter System

Table of content

1 Specifications and basic facts

1.1 Basic facts

The tables below contain mainly specifications by the manufacturer Nova Biomedical.

Table 1. Basic facts

Name of the measurement system:	Nova StatStrip Glucose and ß-Ketone Hospital Meter System (StatStrip)
Dimensions and weight:	Width: 82,5 mm Depth: 46 mm Height: 153 mm Weight: 360 g
Components of the measurement system:	StatStrip Meter Disposable StatStrip test strips Meter docking station
Measurand:	The glucose concentration in plasma
Sample material:	Whole blood: Capillary, venous, arterial, and neonate.
Tolerated haematocrit interval	20 to 65%
Sample volume:	1,2 µL The sample volume is measured by filling the disposable StatStrip test strip.
Measuring principle:	Modified glucoseoxidase-based amperometric test system with haematocrit and chemical interference corrections. Glucose in the blood sample mixes with reagent on the test strip and produces an electric current to the electrode on the test strip. The strength of the electric current is proportional to the glucose concentration in the blood sample.
Traceability:	The system is calibrated by the manufacturer: Venous and arterial specimens were measured parallel with the StatStrip and the YSI 2300 Stat Plus Analyzer (Yellow Springs Instrument Co., Yellow Springs, OH, USA) as reference. Capillary specimens were measured parallel with the StatStrip and the SureStep Flexx Blood Glucose Meter (LifeScan Inc., Milpitas, CA, USA) as reference. The reference methods are traceable to NIST SRM 917a. The StatStrip makes the measurements of the glucose concentration in whole blood samples but as recommended by an IFCC expert committee and like most glucose meters the corresponding glucose concentration in plasma is reported.

Calibration:	The meter is calibrated by the manufacturer. The meter can be recalibrated by the user versus a local measurement system.
Measuring range:	0,6 to 33,3 mmol/L Results below 0,6 mmol/L are shown as “lo” with two red arrows pointing down and high results above 33,3 mmol/L are shown as “hi” with two red arrows pointing up and a flag ”exceeds measure”.
Linearity:	Linear from 0,6 to 33,3 mmol/L Nova Biomedical provides five linearity check control solutions with which the user can check linearity. See below Table 4 Quality control.
Measurement duration:	6 seconds
Operating conditions:	Temperature range: +15°C to +40°C. Altitude: Up to 4500 meters Relative humidity: Up to 90% (noncondensing)
Electrical power supply:	Battery: Rechargeble Li-polymer 3,7 V, 2000 mAh Life: 6-8 hours in use (approximately 40 tests with barcode scans)/12-24 hours standby The meter needs to have the battery charged or battery replaced regularly. The external power supply is connected to the desk-mount docking station: Input 100-240 V , 50-60 Hz, 0,6 A Output +12 V , 0,85 A The docking station has an extra battery slot for recharging and storage of spare battery.
Recommended regular maintenance:	The meter should have its surface cleaned/disinfected regularly.
Package contents:	Meter, Power supply, Docking/Charging Station and documentation
Necessary equipment not included in the package:	Glucose strips and control solutions

Table 2. Post analytical traceability

Is input of patient identification possible?	Yes, by the key-board, by the bar-code reader or by downloading a list from a computer system (ADT-file).
Is input of operator identification possible?	Yes, by the key-board or by the bar-code reader
Can the instrument be connected to a bar-code reader?	A bar-code reader is integrated in the instrument. One--dimensional barcodes supported: a. Code 39 Extended b. Code 93 c. Code 128 d. Interleaved 2 of 5 e. Codabar Two-dimensional barcodes supported: a. Data Matrix b. Maxi Code c. PDF 417 d. QR Code e. Aztec
Can the instrument be connected to a printer?	Yes, through the docking station
What can be printed?	All measurement information can be printed
Can the instrument be connected to a PC?	Yes, through the docking station.
Can the instrument communicate with LIS (Laboratory Information System)? If yes, is the communication bidirectional?	Yes, the instrument can communicate with a LIS system and the communication can be bidirectional. RJ-45 Ethernet (10 Mbit) Protocol TCP/IP Ethernet Standard POCT1-A Compliant
What is the storage capacity of the instrument and what is stored in the instrument?	1000 patient measurements 200 quality control measurements 4000 operators Information for each measurement stored in instrument : - Measured value - Date and time - User ID - Sample ID - Reagent lot number - Comments
Is it possible to trace/search for measurement results?	Yes, a sorted list is presented to the user on the instrument.

Table 3. Facts about the reagent/test strips/test cassettes

Name of the reagent/test strips/test cassettes:	StatStrip GLU Test strip
Stability in unopened sealed vial:	Approximately 18 months
Stability in opened vial:	6 months (or expiry date)
Package contents:	100 strips

Table 4. Quality control

Electronic self check:	The meter carries out electronic internal checks at all stages of the measurement cycle. These ensure that the hardware and software are operating correctly, the strip reader is reading correctly, the glucose strip is not faulty and that the operator is using the system correctly.
Recommended control materials and volume:	Nova Biomedical can deliver control solutions to be used by the StatStrip meter: 1. Three levels of Nova QC Glucose and Ketone Control Solutions: Level 1, Level 2, and Level 3 2. Five levels of Nova Glucose and Ketone Linearity Solutions (values for the full reportable range of meter linearity): Levels 1, 2, 3, 4, and 5 Regular QC check is recommended
Stability in unopened sealed vial:	Approximately 18 months
Stability in opened vial:	3 months (or expiry date)
Package contents:	4 mL

Table 5. Marketing information for the StatStrip system

Manufacturer:	Nova Biomedical Corporation Corporate Headquarter: Nova Biomedical 200 Prospect Street Waltham, MA 02454-9141 U.S.A. Tel: 781-894-0800 Fax: 781-894-5915 Toll Free: 800-458-5813 International Fax: 781-899-0417 Technical Support: 1-800-545-NOVA (6682) Email: info@novabio.com
Retailer in the Nordic Countries:	Sweden: A. Menarini Diagnostics Nordic Countries Medeon Science Park Albin Hanssons Väg 41, hus D 214 32 Malmö Phone: +46 (0)40-32 12 70 Fax: +46 (0)40-32 12 71 Norway: Med-Nett A/S Terassveien 33 B 1363 Høvik Phone: +47 67 82 90 00 Denmark: –
In which countries is the system marketed?:	Globally
Date for start of marketing the system in Scandinavia:	2009-10-01
Date for CE-marking:	May 2007
In which Scandinavian languages is the manual available?:	Norwegian and Swedish

1.2 Imprecision specifications

Nova Biomedical Corporation specifies same values for the typical within-series-imprecision and the total imprecision. See Table 6 below.

Table 6. Imprecision specifications for StatStrip

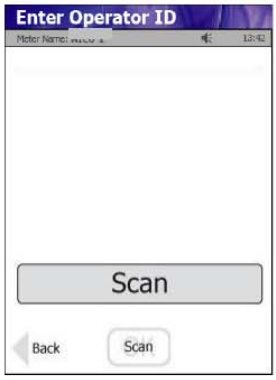
Level	P—Glucose (mmol/L)	Within-series-imprecision and total imprecision (CV %)
1	2,8	8
2	8,3	6
3	22,2	4
4	33,3	4

2 Operating StatStrip

The Quick Operating Guide for StatStrip manual is inserted in this document to make it easier for the reader of the SKUP report to understand how the measurements are performed on the StatStrip system. See next page. These instructions were followed during the present evaluation.

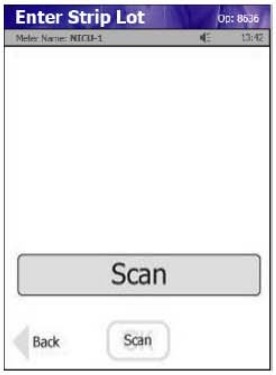
Glucose Monitoring System Quick Operating Guide

1 From Home screen press **Login**.

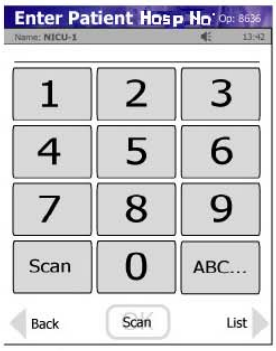


2 Enter or scan Operator ID and press **Accept**.

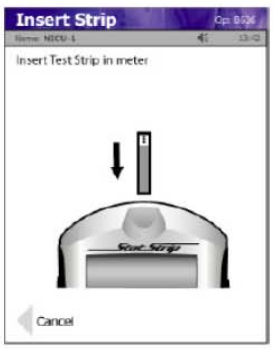
3 From Patient Test screen press **Accept**.



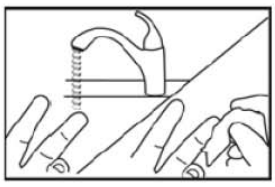
4 Enter or scan Strip lot no. and press **Accept**.



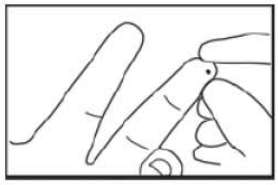
5 Enter or scan Patient ID and press **Accept**. Patient ID must be used, case number or name and DOB are mandatory.



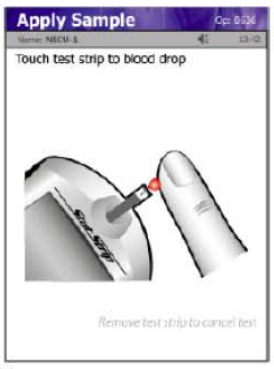
6 Insert Test Strip into Meter.



7 Wash patient's hand thoroughly and massage finger to stimulate blood flow.



8 Use safety lancet to puncture finger. Squeeze finger to form blood drop.



9 Touch strip to blood drop. Result will appear in 6 seconds.

Warning!

The test strip must fill completely upon touching the blood droplet. ***If the test strip does not fill completely, do not touch the blood droplet a second time.*** *Discard the test strip and repeat the test with a new test strip.*

10 *Warning!*

11 To accept result, press **Accept**. To reject result, press **Reject**.

12 To review other results, Press **Review** from Patient Test screen. Follow steps 1 and 2 first.

nova biomedical

Nova Biomedical, C3 - 5 Evans Business Centre, Deeside Industrial Park, Deeside, Flintshire, CH5 2JZ
Tel: 01244 287087 • www.novabiomedical.com

Rev 2.0 19th December 2007

3 Agreement between StatStrip meters

3.1 Check of the calibration agreement between meters

Agreement between the StatStrip meters used in the present evaluation was checked before the evaluation. The four meters were placed next to each other in the hospital laboratory. Two patient samples, one with low and one with high P—Glucose concentration were selected from the routine blood gas samples. Both samples were measured six times on each StatStrip meter.

To avoid influence from glycolysis during the measurements, the measurement was done according to the order number in the Table 7 below. Explanation: The first measurement was done with Meter 1, the second with Meter 2, ..., the fourth and the fifth with Meter 4 and so on according to the table. All 24 measurements were done in one sequence aiming at equal time difference between the measurements. That time difference corresponded to something between five and ten seconds. However the assigned time points were in Table 7 used to keep the "mean time" of all measurements performed with one instrument and achieve the same"time difference" between the first and the last measurement with the four instruments.

Table 7. Order of measurements when checking agreement between StatStrip meters

	Meter 1	Meter 2	Meter 3	Meter 4
Measurement 1	1	2	3	4
Measurement 2	8	7	6	5
Measurement 3	12	11	10	9
Measurement 4	13	14	15	16
Measurement 5	20	19	18	17
Measurement 6	21	22	23	24
Mean time:	12,5	12,5	12,5	12,5
Time difference:	20	20	20	20

SKUP has in earlier evaluations used the requirement that the CV is allowed to be 30% higher as a maximum, when the total imprecision of the results from all meters is compared with the mean imprecision within several individual meters. If the total CV is higher, the meters do not fulfil the required conformity. In that case, the meter(s) with deviating mean value or deviating CV should be identified and excluded from the evaluation. The manufacturer should, in such a case, be contacted for exchange of the deviating meter. This model for assessing the conformity has been used also in this evaluation.

The results of the agreement check are shown in the Tables 8 and 9 below.

Table 8. Agreement between different StatStrip meters

	Meter serial number			
	140099 110242	140035 310168	140098 910242	140000 910209
Evaluation site	Hospital laboratory	Neonatal ward	Maternity ward	Endocrine clinic
Sample no. 1 (n = 6)				
P—Glucose mean (mmol/L)	5,08	4,88	4,92	4,97
CV (%)	4,2	4,0	4,7	4,0
Sample no. 2 (n = 6)				
P—Glucose mean (mmol/L)	10,13	9,95	10,08	10,00
CV (%)	4,4	3,3	2,3	2,8
Both samples				
P—Glucose mean (mmol/L)	7,6	7,4	7,5	7,5
CV mean (%)	4,3	3,6	3,5	3,4

As can be seen in Table 8 the means and CVs of all instruments agreed well.

Table 8 also shows which instrument that was used at each site. There were negligible calibration differences between the StatStrip instruments at the different evaluation sites.

Table 9. ANOVA calculations of agreement between different StatStrip instruments

Sample no.	P—Glucose mean (mmol/L)	Within-instrument CV component (%)	Between-instruments CV component (%)	Total CV (%)	Increase of CV (%)
1	4,96	4,2	0,4	4,2	0,4
2	10,04	3,3	0,0	3,1	0,0

"Within-instrument CV component" refers to the mean contribution to the "Total CV" originating from the within instrument imprecision.
"Between-instruments CV component" refers to the mean contribution to the "Total CV" originating from the between instruments imprecision. The within-instrument imprecision is not included in this figure.

"Increase of CV" refers to the increase in percent of the CV from "Within-instrument CV component" to "Total CV".

As can be seen in Table 9 the values for “increase of CV” were less than 1%.
The requirements defined by SKUP for agreement between instruments, “increase of CV” less than 30%, were fulfilled by a large margin.

A. Menarini was informed of these results as soon as they were ready. A. Menarini accepted the decision to use all the tested instruments for the evaluation.

3.1.1 Conclusion of the check of the agreement between meters

All four meters tested before the evaluation showed good calibration agreement and the variation of the results increased very little by using several meters. All four instruments were accepted to be used in the evaluation.

4 Internal quality control results with StatStrip

4.1 Intermediate imprecision with control solutions in the hospital laboratory

The daily internal quality control results were used for calculation of the intermediate imprecision in Table 10.

Table 10. Internal quality control results with control solutions at the hospital laboratory

StatStrip control	P—Glucose assigned value (interval)[1] (mmol/L)	n	Number of excluded results	StatStrip mean (min. — max.) (mmol/L)	SD[2] (mmol/L)	CV[2] (95% confidence interval) (%)
Low	3,4 (2,6 — 4,2)	14	0	3,4 (3,3 — 3,6)	0,14	4,1 (2,9 — 6,5)
High	16,4 (14,4 — 18,3)	14	0	16,6 (15,9 — 17,5)	0,59	3,6 (2,6 — 5,8)

[1] Assigned values and intervals for acceptance are set by the manufacturer.
[2] In addition to the pure repeatability imprecision, the calculated SD and CV values include some variance components arising from changes in conditions during the collection of measurement data: three different batches of test strips were used and the measurements were performed different days.

4.2 Intermediate imprecision with control solutions at the hospital wards

The daily internal quality control results were used for calculation of the intermediate imprecision in Table 11.

Table 11. Intermediate imprecision of StatStrip with internal quality control solutions at the hospital wards

StatStrip control	P—Glucose assigned value (interval)[1] (mmol/L)	n	Number of excluded results	StatStrip mean (min. — max.) (mmol/L)	SD[2] (mmol/L)	CV[2] (95% confidence interval) (%)
Maternity Ward:						
Low	3,4 (2,6 — 4,2)	7	0	3,3 (3,2 — 3,6)	0,16	4,8 (3,1 — 10,7)
High	16,4 (14,4 — 18,3)	7	0	16,8 (16,4 — 17,1)	0,25	1,5 (1,0 — 3,3)
Endocrine clinic:						
Low	3,4 (2,6 — 4,2)	5	0	3,4 (3,2 — 3,6)	0,19	5,5 (3,3 — 15,8)
High	16,4 (14,4 — 18,3)	5	0	16,4 (15,3 — 17,4)	0,98	6,0 (3,6 — 17,2)

[1] Assigned values and intervals for acceptance are set by the manufacturer.
[2] In addition to the pure repeatability imprecision, the calculated SD and CV values include some variance components arising from changes in conditions during the collection of measurement data: and three different batches of test strips were used and the measurements were performed varying days.

5 Additional equipment and product details

5.1.1 StatStrip meters and software

Four StatStrip meters were originally available for the evaluation, but only three were used. The serial numbers of used instruments used at each evaluation site are shown in Table 8.

The StatStrip meters with serial numbers 140000910209, 140099110242 and 140098910242 were used in the evaluation. The software in these StatStrip meters had the following version numbers: Host version 1.1.8.5, OS version 1.0.5.3 and CF version 2.0.7045.0

5.1.2 StatStrip test strips

At the check of agreement between different StatStrip meters before the evaluation the used test strips had the lot number 0310193249.

Three different lots of test strips were used in the evaluation:
Lot 310144249 with expiry date 2012-05-31
Lot 310223249 with expiry date 2012-08-31 and
Lot 310193249 with expiry date 2012-07-31

Approximately one third of the measurements were performed with each lot at each evaluation site. The two measurements in each duplicate were always performed with the same lot.

5.1.3 Material for internal quality control

The control solutions used for StatStrip during the evaluation was supplied by
A. Menarini Diagnostics and manufactured by Nova Biomedical:

Nova QC Glucose and Ketone Control Solutions:
Level 1, Lot 0410154301, expiry date 2012-12
Level 3, Lot 0410154303, expiry date 2012-12

5.1.4 Syringes for collection of arterial blood samples

The arterial blood samples in the hospital laboratory evaluation were collected in the following type of syringe:
Portex Line Draw Plus Dry 3cc Syringe, Luer Slip, Filter-Pro. Product code: 4043E.
The syringe contains balanced dry lithium heparin and when filled with 3 mL blood the heparin concentration will be 23,5 U/mL
Supplied by Smiths Medical Sverige AB, Sweden, E-mail: info.sweden@smiths-medical.com

5.1.5 Lancets for capillary punctures

Capillary punctures were made only at the Endocrine Clinic. They used the following lancets: Prolance®, Normal Flow (green), needle style safety lancets, needle diameter 21G (gauge), penetration depth 1,8 mm, product number 7594, supplied by MedCore Sweden AB, Sweden, E-mail: info@medcore.se

5.1.6 Micro tubes for capillary blood collection

The capillary samples for the comparison method were collected in the following micro tubes:
Microvette 300, additive: lithium heparin plasma, product number 20.1309,
supplied by Sarstedt AB, Sweden, E-mail: info@sarstedt.com

5.1.7 Injection needles for venous sampling

For the venous sampling from the newborn children in the Maternity ward the following injection needles were used:
Sterile injection needles, 0,6 x 25 mm, 23 G x 1", brand: KD-Fine, subsupplier: MediCarrier, Sweden, their product number: 63741, agent for Sweden: OneMed Sverige AB, their product number: 295703, E-mail: kundservice@onemed.com

5.1.8 Micro tubes for venous sampling

For the venous sampling from the newborn children in the Maternity ward the following micro tubes were used:
Microvette 300, additive: lithium heparin plasma, product number 20.1309,
supplied by Sarstedt AB, Sweden, E-mail: info@sarstedt.com

5.1.9 Micro tubes for separated plasma

The blood samples for the comparison method was centrifuged and the separated plasma was transferred to the following micro tubes:
Micro tubes, 2,0 mL, Type I, with skirted base, neutral screw cap, product number 72.694, supplied by Sarstedt AB, Sweden, E-mail: info@sarstedt.com

Attachment 2.
The selected comparison method

Table of content

The selected comparison method

The selected comparison method in this evaluation was the routine method for P—Glucose in the Department for Clinical Chemistry at Karolinska University Hospital in Huddinge. It is a Roche hexokinase method, Gluco-quant Glukos/HK, applied on a Modular Analytics P instrument from Roche Diagnostics. This method is put in practice completely according to the instructions from Roche. It is here on called "the comparison method".

1 Description of the comparison method

1.1 Facts about the comparison method

1.1.1 Measurement principle

The method is based on the following reactions

$$\text{Glucose} + \text{ATP} \xrightarrow{\text{Hexokinase}} \text{G-6-P} + \text{ADP}$$

In the presence of ATP, hexokinase catalyses the phosphorylation of glucose to glucose-6-phosphate.

$$\text{G-6-P} + \text{NADP}^+ \xrightarrow{\text{Glucose-6-phosphate dehydrogenase}} \text{gluconate-6-P} + \text{NADPH} + \text{H}^+$$

In the presence of NADP, glucose-6-phosphate dehydrogenase oxidises glucose-6-phosphate to gluconate-6-phosphate. No other carbohydrate is oxidised.

The speed of the formation of NADPH during the reaction is proportional to the glucose concentration. The speed of NADPH formation is measured photometrically.

1.1.2 Instrument

Modular Analytics P instrument from Roche Diagnostics.

1.1.3 Reagents

R1 TRIS-buffer: 100 mmol/L, pH 7,8; Mg^{2+}: 4 mmol/L; ATP ≥1,7 mmol/L; NADP ≥ 1,0 mmol/L; preservative
R2 reaction starter: HEPES-buffer: 30 mmol/L, pH 7,0; Mg^{2+}: 4 mmol/L; HK (yeast) ≥8,3 U/mL; G-6-PDH (E. coli) ≥15 U/mL; preservative.
All reagents are supplied by Roche.

1.1.4 Calibration

Calibrator: C.f.a.s. (Calibrator for automated systems), cat.no:10759350 190
Routine two-point calibration with S1: 0,9 % NaCl and S2: C.f.a.s.
Calibration frequency:
• after change to a new lot of reagent
• if required according to internal quality control results outside set limits

1.1.4.1 Traceability

This method is standardised to a ID-MS-method.

1.1.5 Measuring range

P—Glucose: 0,11 to 41,6 mmol/L

Samples with higher concentrations are measured after dilution.

1.1.6 Additional equipment and product details

Additional equipment and product details for equipment used in the evaluation are specified in Attachment 1, Section 5

2 Verification of the comparison method

2.1 Imprecision of the comparison method

2.1.1 Missing and excluded results and check calculations

See Attachment 4, Section 3, Table 1B.

2.1.2 Imprecision of the comparison method in the StatStrip evaluation

For each patient in the hospital laboratory evaluation the same tube with arterial sample was first used for measurements with StatStrip and the comparison method. After measurements with StatStrip, the sample was then centrifuged and plasma separated and used for measurements with the comparison method. Both methods measured in duplicates.

A requirement for correct imprecision calculations is that the results in each calculation show homogenous or similar variation (homoscedasticity). To check the homogeneity of the variation in the results the diagram in Figure 1 was drawn. In the diagram it is obvious that the frequency of high $(d/m)^2$ values is high at low concentration, below 2,5 mmol/L. These results don't show homoscedasticity compared to the rest of the results and the imprecision of these results are therefore calculated in a separate level group.

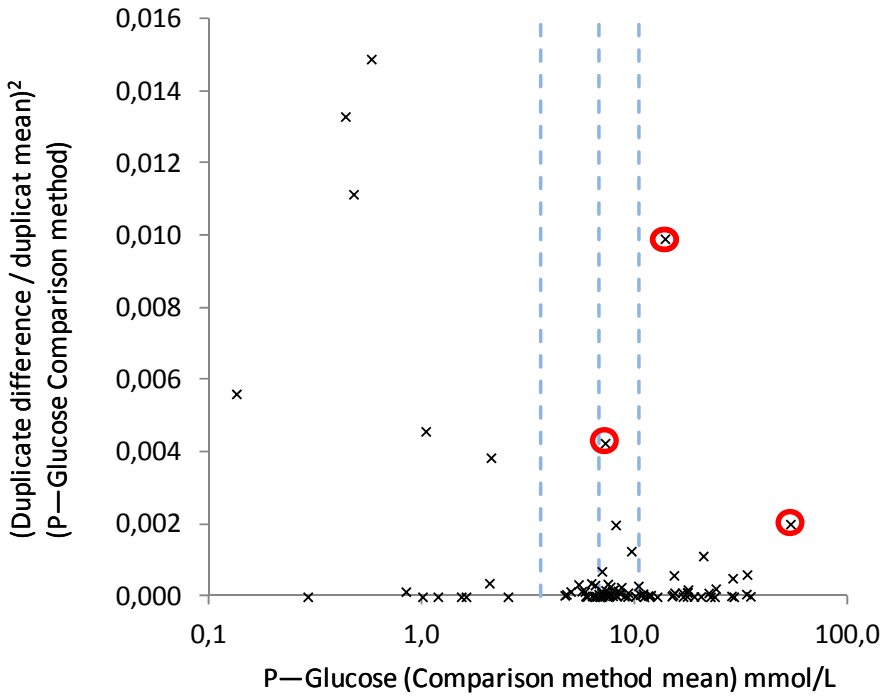


Figure 1. Imprecision profile for the comparison method resuts
The x axis shows the mean glucose concentration of each duplicate. The y axis shows the $(d/m)^2$ values included in the imprecision calculations. Each cross represents a duplicate result. The limits for the different level groups are drawn as vertical stippled blue lines. The red rings mark duplicate results which are identified and excluded by the outlier test.

For each patient in the evaluation at the Maternity Ward the same tube with venous sample was used for measurements with StatStrip and the comparison method. After measurements with StatStrip, the sample was then centrifuged and plasma separated and used for measurements with the comparison method. Both methods measured in duplicates.

For each patient in the evaluation at the Endocrine Clinic, duplicate capillary measurements were first performed with StatStrip. A capillary sample was then collected for measurements with the comparison method. The measurements with the comparison method were performed in duplicates.

The repeatability of the comparison method within each level group was calculated from the duplicate comparison method results from each evaluation site. The outlier test was performed on relative differences (d/m) and formula 2 described in Attachment 3, section 2.2 was used for the calculation. See table 1.

Table 1. Repeatability of the comparison method.

Level[#]	Comparison method interval (mmol/L)[#]	n	Number of excluded results*	Comparison method mean (mmol/L)[#]	CV (%) (95 % confidence interval)
Hospital laboratory:					
Very low	0,13 — 2,5	14	0	1,1	4,4 (3,2 — 7,1)
Low	4,7 — 6,8	19	0	6,0	0,7 (0,5 — 1,0)
Medium	6,9 — 10,4	32	1	8,0	1,0 (0,8 — 1,4)
High	10,6 — 34,9	30	2	19,6	0,8 (0,7 — 1,1)
Maternity Ward:					
Low	3,1 — 4,5	21	0	4,0	1,3 (1,0 — 1,9)
Medium	4,5 — 6,4	20	0	5,0	0,8 (0,6 — 1,2)
Endocrine Clinic:					
Medium	4,9 — 10,0	20	0	7,1	0,8 (0,6 — 1,2)
High	10,0 — 27,2	20	0	15,3	0,7 (0,6 — 1,1)

[#] The results are divided into concentration subgroups according to the comparison method results. The groups contain the same sample results as the results in Table 4 and Table 6 in the StatStrip report. So the tables are direct comparable

* n is the number of results before exclusion of outliers.

2.1.3 Internal quality control results of the comparison method

The following internal quality control materials are used to check the comparison method:
Bio-Rad Liquid Unassayed Multiqual, level 1, article number: 697, lot number 46401
Bio-Rad Liquid Unassayed Multiqual, level 3, article number: 699, lot number 46403

Table 2 contains the internal quality control results obtained with the comparison method during 2011. The target value is assigned after two weeks use of new material. An assigned value is sometimes changed to fit reality. Assigned values during 2011:
For level 1: During the period 2011-01-01 to 2011-06-15 3,4 mmol/L. For the rest of the year 3,3 mmol/L.
For level 3: During the period 2011-01-01 to 2011-01-31 20,0 mmol/L. For the rest of the year 19,5 mmol/L.

Table 2. Internal quality control results of the unadjusted comparison method for 2011

Quality control level	P—Glucose, comparison method (mmol/L)		n	Accepted SD (mmol/L)	Found SD (mmol/L)	Found CV (%)
	Assigned value	Found average				
Level 1	3,3 to 3,4	3,335	658	0,100	0,076	2,29
Level 3	19,5 to 20,0	19,58	1968	0,500	0,484	2,47

2.1.4 Assessment of the imprecision of the comparison method

The CV, calculated from the duplicate measurements on patient samples in the evaluation was about 1%. The CV for the internal quality control results was maximum 2,5 %.
The imprecision figures of the comparison method are considered to be good and normal for a hospital laboratory method.

2.2 Trueness of the comparison method

2.2.1 Results of the unadjusted comparison method in EQA

The comparison method show normal P—Glucose results compared to other hospital laboratory methods in the Equalis EQA scheme for general clinical chemistry in Sweden. See table 3. The used sample materials were pooled human sera.

Table 3. EQA results for the unadjusted comparison method in the Swedish EQA scheme for general clinical chemistry

Round (year-week no)	P--Glucose, comparison method (mmol/L)	Number of laboratories with the same method	P--Glucose, method mean (mmol/L)	P--Glucose, method group SD (mmol/L)	Bias from method group mean (number of SD)	Bias from method group mean (%)	Number of laboratories, totally	P--Glucose, total mean (mmol/L)	P--Glucose, total EQA SD (mmol/L)	Bias from total EQA mean (number of SD)	Bias from total EQA mean (%)
Pooled unmodified serum:											
2011-02	3,89	33	3,86	0,0785	+0,4	+0,9	91	3,88	0,1083	+0,2	+0,5
2011-10	4,52	33	4,32	0,0735	+2,8	+4,7	89	4,33	0,1096	+1,7	+4,4
2011-20	4,43	34	4,32	0,0475	+2,4	+2,7	87	4,32	0,0918	+1,3	+2,7
2011-34	3,80	35	3,85	0,0658	−0,7	−1,2	88	3,85	0,0916	−0,5	−1,3
2011-45	3,87	37	3,85	0,0549	+0,4	+0,5	94	3,84	0,0972	+0,3	+0,7
Mean	4,10					+1,5					+1,4
Modified serum:											
2011-06	3,35	31	3,29	0,0557	+1,2	+2,0	88	3,28	0,1165	+0,6	+2,2
2011-15	15,6	34	15,4	0,2883	+0,5	+0,9	90	15,5	0,3265	+0,2	+0,4
2011-24	3,33	33	3,30	0,0644	+0,5	+1,0	81	3,29	0,1001	+0,5	+1,4
2011-39	8,27	36	8,35	0,1315	−0,6	−1,0	89	8,37	0,1629	−0,6	−1,2
2011-50	15,4	38	15,4	0,2458	−0,1	−0,2	95	15,5	0,2816	−0,5	−0,9
Mean	9,19					+0,5					+0,4

2.2.2 The trueness of the unadjusted comparison method in EQA

The trueness of the unadjusted comparison method was good judged from EQA results. The trueness is similar to that of other hospital laboratory methods in Sweden.

2.2.3 Calibration adjustment of the comparison method

The samples in the evaluation were measured with the comparison method in two series. 2011-02-10 the 18 first samples in the hospital laboratory evaluation and 2011-11-17 the remaining samples.

The calibration of the comparison method was checked with the SRM 965b from NIST in both measurement series of the evaluation. The agreement between the comparison method and the NIST-standards is shown in table 4 and 5.

Table 4. Standard Reference Material (SRM 965b) measured with the unadjusted comparison method in the first series of measurements in the evaluation 2011-02-10

	Certified glucose concentration, mmol/L (uncertainty)	n	Measured mean glucose (mmol/L)	% deviation from target value
Level 1	1,836 (1,809 to 1,863)	5	1,832	−0,22
Level 2	4,194 (4,135 to 4,253)	5	4,278	+2,00
Level 3	6,575 (6,481 to 6,669)	5	6,822	+3,76
Level 4	16,35 (16,15 to 16,55)	5	16,702	+2,15

Table 5. Standard Reference Material (SRM 965b) measured with the unadjusted comparison method in the second series of measurements in the evaluation 2011-11-17

	Certified glucose concentration, mmol/L (uncertainty)	n	Measured mean glucose (mmol/L)	% deviation from target value
Level 1	1,836 (1,809 to 1,863)	5	1,832	−0,22
Level 2	4,194 (4,135 to 4,253)	5	4,218	+0,57
Level 3	6,575 (6,481 to 6,669)	5	6,644	+1,05
Level 4	16,35 (16,15 to 16,55)	5	16,270	−0,49

In table 4 the measured glucose results of the NIST-standards on level 2, 3 and 4 were above the upper uncertainty limits. In table 5 all measured glucose results were inside the uncertainty limits. All results from the comparison method were adjusted according to the certified NIST-targets. The adjustment was carried out by means of ordinary linear regression by the following adjustment equations:
for the short series run 2011-02-10: Adjusted value = 0,9775 × Unadjusted value − 0,0029
and
for the long series run 2011-11-17: Adjusted value = 1,0068 × Unadjusted value − 0,0517

Approximately 90% of the comparison method results in the evaluation were measured in the second series, for which the adjustment was negligible. Further on in the report, whenever any result from the comparison method is presented, the result has already been adjusted according to the above equations.

2.2.4 The trueness of the calibration adjusted comparison method

To verify the trueness of the comparison method, EQA materials, human serum controls used in Equalis EQA scheme for general clinical chemistry, were measured. The agreement between the comparison method and the target value from a reference laboratory is shown in table 6.

Table 6. Trueness of the adjusted comparison method

Material used in EQA year-week number	Type of sample material and responsible reference laboratory	P--Glucose, reference method (Uncertainty, k = 2) (mmol/L)	P--Glucose, EQA total mean (mmol/L)	P--Glucose, EQA Roche method mean (mmol/L)	P--Glucose, comparison method (mmol/L) 2011-11-17	Difference from reference method value (mmol/L)	Difference from reference method value (%)
Several	UL	8,21 (8,09—8,30)	8,42	8,39	8,25	+0,04	+0,5
2009-50	MD	4,31	4,27	4,27	4,07	−0,24	−5,6
2010-24	MD	11,0	10,7	10,6	10,56	−0,44	−4,0
2010-45	UL	3,58 (3,55—3,61)	3,64	3,64	3,37	−0,21	−5,8
2011-15	ML	15,06	15,5	15,4	15,21	+0,15	+1,0

Type of sample material: U = unmodified pooled human serum spiked with natural serum components only, M = modified pooled human serum, spiked with animal enzymes and drugs.
Responsible reference laboratory: L = Ulf Hannestad, Linköping, D = DGKL

The first material in table 6 has been determined with reference method at three occasions during the years 2005 to 2011 with four measurements each time. The determinations produced the following mean results 8,17, 8,41 and 8,05 – on average 8,21 mmol/L.
The "EQA total mean" and "EQA Roche" for this material are mean values from three rounds in the Equalis EQA program for general clinical chemistry in Sweden. The comparison method result is within the uncertainty of the reference method value so with this material there was no bias.

The material 2010-45 shows a large negative bias, which we can't explain. The narrow uncertainty for this material is based on within-series imprecision only and is probably false too low.

The varying results with modified sera are probably explained by sample specific influences on the results, i.e. matrix effects. We know that modified sera which are spiked with animal enzymes and drugs produce different bias than unmodified human samples. So those results should be trusted less.

2.2.5 Assessment of the trueness of the comparison method

The comparison method results showed a negative bias with one material determined with a reference method once and good trueness with another material determined with a reference method at three occasions.

Attachment 3.
Statistical expressions and calculations

Table of content

This attachment with standardised text deals with the statistical terms, expressions and calculations used by SKUP. The attachment is a short extract of the comprehensive SKUP-document "Statistics in SKUP reports", available at the SKUP website [1]. The statistical calculations will change according to the type of evaluation. The descriptions in this section are valid for evaluations of quantitative methods with results on the ratio scale.

1 Statistical terms and expressions

The definitions in this section come from the ISO/IEC Guide 99; International Vocabulary of Metrology, VIM [2].

1.1 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, intermediate, 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 unit 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.

1.2 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, intermediate, poor e.g.), whereas the bias is reported in the same unit as the analytical result or in percent.

1.3 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, intermediate, poor e.g.).

2 Statistical calculations

2.1 Statistical outliers

The criterion promoted by Burnett [3] 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.

2.2 Calculation of imprecision

The precision of the field method is assessed by use of paired measurements of genuine patient sample material. The results are divided into three concentration levels, and the estimate of imprecision is calculated for each level separately, using one of the following formulas:

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

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

(Formula 1)

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 [4, 5] 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 assumption for using the formulas is that there is no systematic difference between the 1st and the 2nd measurement of the pairs.

2.3 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 field method. The mean difference is shown with a 95% confidence interval.

2.4 Assessment of accuracy

The agreement between the field 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 field 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.

3 References

1. www.skup.nu / ”The SKUP evaluation” (in the left menu) / “Statistics and calculations”/”Statistics in the SKUP reports, version 1.0”.
2. ISO/IEC Guide 99:2007, International vocabulary of metrology – Basic and general concepts and associated terms, VIM, 3rd edition, JCGM 200:2008
3. Burnett RW, “Accurate Estimation of Standard Deviations for Quantitative Methods Used in Clinical Chemistry”. Clinical Chemistry 1975; 21 (13): 1935 – 1938
4. 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”, ISBN 0-7216-0189-8
5. Fraser, C.G, Biological variation: From principles to practice. 2006. Chapter 1 “The Nature of Biological Variation”. AACC Press. ISBN 1-890883-49-2

Attachment 4.
Missing or excluded results and check calculations

Table of content

1 Explanation of the content in this attachment

For some samples some results are missing. For other samples some results are excluded as statistical outliers. All missing or excluded results are explained in this attachment to show that the raw data has been treated correct and that exclusion of data has been done in a consequent manner.

How the calculations of imprecision from duplicates have been done is explained in chapter 3.7 in the main report about StatSrip and in the more comprehensive document ”Statistics in the SKUP reports, version 1.0” available at the SKUP website [1]. The used formula will not produce correct CV values if there is a systematic difference between the results of the first and the second measurements. All the result groups are therefore tested for such differences. The tables in this attachment show the mean differences with confidence intervals. There is no systematic difference between the first and the second measurements in the duplicates if the confidence interval of the difference includes zero or includes numbers very close to zero. The conclusion is thus that the CV calculations are valid.

The numbering of the tables in this attachment follows the numbering of the tables in the main StatStrip report.

2 Missing and excluded results

2.1 Missing and excluded results in the hospital laboratory evaluation

There were a total of 102 patient results. A detailed explanation of applied tests for exclusion of statistical outliers is given in the document ”Statistics in the SKUP reports, version 1.0” [1]. The explanation is found in Section 4. Which and why some results are missing or excluded are shown in Table 4A below. The number of results remaining in respective calculation and in the diagram is shown in the table.

Table 4A. Missing or excluded results in the hospital laboratory evaluation

Which results: Explanation of why missing or excluded	Calculation of the comparison method repeatability	Calculation of the StatStrip repeatability	Calculation of the StatStrip bias	Diagram showing influence of haematocrit	Difference plot	Calculation of the percentage of results within allowable limits
All results:	102	102	102	102	102	102
Missing results:						
Result #47: Duplicate comparison method result is missing	-1	-1	-1	-1	-1	-1
Result #79: Duplicate StatStrip result is missing	-1	-1	-1	-1	-1	-1
Results #68, #69, #78 and #85: Haematocrit results are missing	–	–	–	-4	–	–
Excluded incalculable results:						
Results #73 and # 84: Both comparison method duplicate results are below the measuring range of the comparison method and both StatStrip duplicat results are below the measuring range of StatStrip	-2	-2	-2	-2	-2	–
Results #76, #83 and #89: Both StatStrip duplicate results are below the measuring range of StatStrip	–	-3	-3	-3	-3	–
Results #74 and #75: The first of theStatStrip duplicate results are below the measuring range of StatStrip	–	-2	-2	-2	-2	–
Result #90: The second of the StatStrip duplicate results is below the measuring range of StatStrip	–	-1	-1	-1	–	–
Results #77 and #81: Both StatStrip duplicate results are above the measuring range of StatStrip	–	-2	-2	-2	-2	–
Result #82: The first StatStrip duplicate result is above the measuring range of StatStrip	–	-1	-1	-1	-1	–
Result #64: The second StatStrip duplicate result is above the measuring range of StatStrip	–	-1	-1	-1	–	–
Excluded outliers:						
Results #28, #77 and #88: Outliers with deviating high differences between the two comparison method duplicate results. However #77 is in some columns already excluded on a previous line in the table.	-3	–	-2	-2	-2	-3
Results with deviating high differences between the two StatStrip method duplicate results	0	0	0	0	–	–
Results with deviating high differences between the StatStrip mean results and the comparison method mean result	–	–	0	0	–	–
Number of results included:	95	88	86	82	88	97

"–" means that these results should not be excluded

"-1" means that one result is missing/excluded

"0" means that there were no results of this category

In cells with blue background result #77 is not counted as it is already excluded on a previous line in the table

2.2 Missing or excluded results in the evaluations in the hospital wards

Totally there were 41 + 40 patient results. A detailed explanation of applied tests for exclusion of statistical outliers is given in the document ”Statistics in the SKUP reports, version 1.0” [1]. The explanation is found in Section 4. Which and why some results are missing or excluded are shown in Table 6A and 6B below. The number of results remaining in respective calculation and in the diagram is shown in the tables.

Table 6A. Missing or excluded results in the Maternity ward evaluation

Which results: Explanation of why missing or excluded	**Calculation of the comparison method repeatability**	**Calculation of the StatStrip r epeatability**	**Calculation of the StatStrip bias**	**Difference plot**	**Calculation of the percentage of results within allowable limits**
All results:	41	41	41	41	41
Missing results:	0	0	0	0	0
Excluded incalculable results:					
Results outside the measuring interval of the comparison method	0	–	0	0	–
Results outside the measuring interval of StatStrip	–	0	0	0	–
Excluded outliers:					
Outliers with deviating high differences between the two comparison method duplicate results.	0	–	0	0	0
Outliers with deviating high differences between the two StatStrip duplicate results	–	0	0	–	–
Result #4334: Result with deviating high difference between the StatStrip mean result and the comparison method mean result	–	–	-1	–	–
Number of included results:	41	41	40	41	41

"–" means that these results should not be excluded

"-1" means that one result is missing/excluded

"0" means that there were no results of this category

Table 6B. Missing or excluded results in the Endocrine clinic evaluation

Which results: Explanation of why missing or excluded	Calculation of the comparison method repeatability	Calculation of the StatStrip r epeatability	Calculation of the StatStrip bias	Difference plot	Calculation of the percentage of results within allowable limits
All results:	40	40	40	40	40
Missing results:					
Sample #26: Duplicate measurement on StatStrip is not complete. The second value is missing.	–	-1	-1	–	–
Excluded incalculable results:					
Results outside the measuring interval of the comparison method	0	–	0	0	–
Results outside the measuring interval of StatStrip	–	0	0	0	–
Excluded outliers:					
Outliers with deviating high differences between the two comparison method duplicate results.	0	–	0	0	0
Outliers with deviating high differences between the two StatStrip method duplicate results	–	0	0	–	–
Outliers with deviating high difference between the StatStrip mean result and the comparison method mean result	–	–	0	–	–
Number of results included:	40	39	39	40	40

"–" means that these results should not be excluded

"-1" means that one result is missing/excluded

"0" means that there were no results of this category

3 Check of the imprecision calculations

3.1 Check of the repeatability calculations for the comparison method

The calculation of imprecision of the comparison method is presented in Table 1 in Attachment 2 to the StatStrip report. The Table 1A below shows the check of differences between first and the second measurements in the same results group. The confidence intervals of the differences for the result groups include or are very close to 0,00 mmol/L. For the total set of data the conclusion is that there is no systematic difference between the first and the second measurements in the duplicates. The calculated CV values in Table 1 are thus valid.

Table 1A. Differences between the 1st and the 2nd measurements with the comparison method

Level	**1. – 2. mean difference (95 % confidence interval)** (mmol/L)
Results in the hospital laboratory evaluation:	
Very low	+0,03 (+0,01 — +0,05)
Low	+0,02 (−0,01 — +0,04)
Medium	±0,00 (−0,04 — +0,04)
High	−0,12 (−0,22 — −0,03)
Results in the maternity ward evaluation:	
Low	−0,02 (−0,05 — +0,01)
Medium	−0,03 (−0,05 — −0,01)
Results in the endocrine clinic evaluation:	
Medium	−0,06 (−0,09 — −0,03)
High	−0,08 (−0,14 — −0,03)

3.2 Check of the repeatability calculations for StatStrip

The calculations of repeatability for StatStrip with patient samples in the evaluation are presented in Table 4 and Table 6 in the main report. The Table 4B below shows the check of the mean differences between first and the second measurements in the same result group. The confidence intervals of the mean differences include 0,00 mmol/L for all result groups except one. For these sets of data, the conclusion is that there is no systematic difference between the first and the second measurements in the duplicates. The calculated CV values in Table 4 and Table 6 are thus valid except for the results commented below.

For the high results in the Endocrine clinic, the mean of the second duplicate results is higher than the mean of the first results. The confidence interval for the mean difference does not include zero difference. If CV is calculated from such results, the CV is not a "true imprecision CV". The consequence of this in Table 6 in the report was that no CV was calculated from these results.

Table 4B. Differences between the 1st and the 2nd measurements with StatStrip

Level	**StatStrip 1. – 2. mean difference (95 % confidence interval) (mmol/L)**
Results in the hospital laboratory evaluation:	
Very low	+0,01 (−0,10 — +0,12)
Low	+0,01 (−0,10 — +0,12)
Medium	−0,09 (−0,21 — +0,03)
High	±0,00 (−0,29 — +0,30)
Results in the Maternity ward evaluation:	
Low	+0,09 (−0,08 — +0,27)
Medium	+0,10 (−0,04 — +0,23)
Results in the Endocrine clinic evaluation:	
Medium	±0,00 (−0,31 — +0,31)
High	+0,73 (+0,11 — +1,34)

4 Reference

1 . www.skup.nu Click on "The SKUP evaluation" (in the left menu), then "Statistics and calculations" and then "Statistics in the SKUP reports, version 1.0".

List of previous SKUP evaluations

Summaries and complete reports from the evaluations are found at www.skup.nu. In addition, SKUP reports are published at www.skup.dk, where they are rated according to the national Danish quality demands for near patient instruments used in primary health care. SKUP summaries are translated into Italian by Centre for Metrological Traceability in Laboratory Medicine (CIRME), and published at http://users.unimi.it/cirme. SKUP as an organisation has no responsibility for publications of SKUP results on these two web-sites.

SKUP evaluations from number 60 and further

Evaluation no.	Component	Instrument/testkit	Producer
SKUP/2013/85	Glucose	StatStrip	Nova Biomedical
SKUP/2013/96	Haemoglobin	DiaSpect Hemoglobin T	DiaSpect Medical GmbH
SKUP/2012/95	Glucose[1]	Mendor Discreet	Mendor Oy
SKUP/2012/94	Glucose[1]	Contour XT	Bayer Healthcare
SKUP/2011/93*	Glucose	Accu-Chek Performa	Roche Diagnostics
SKUP/2012/91	HbA1c	Quo-Test A1c	Quoient Diagnostics Ltd
SKUP/2011/90	CRP	*i*-Chroma	BodiTech Med. Inc.
SKUP/2010/89*	Glucose	FreeStyle Lite	Abbott Laboratories
SKUP/2010/88*	HbA1c	*Confidential*	
SKUP/2011/86	Glucose[1]	OneTouch Verio	LifeScan, Johnson & Johnson
SKUP/2011/84*	PT-INR	Simple Simon PT and MixxoCap	Zafena AB
SKUP/2010/83*	Glucose	*Confidential*	
SKUP/2010/82*	Glucose, protein, blood, leukocytes, nitrite	Medi-Test URYXXON Stick 10 urine test strip and URYXXON Relax urine analyser	Macherey-Nagel GmBH & Co. KG
SKUP/2010/81*	Glucose	mylife PURA	Bionime Corporation
SKUP/2010/80	PT (INR)	INRatio2	Alere Inc.
SKUP/2010/79*	Glucose, protein, blood, leukocytes, nitrite	CombiScreen 5SYS Plus urine test strip and CombiScan 100 urine analyser	Analyticon Biotechnologies AG
SKUP/2010/78	HbA1c	In2it	Bio-Rad
SKUP/2011/77	CRP	*Confidential*	
SKUP/2009/76*	HbA1c	*Confidential*	
SKUP/2009/75	Glucose	Contour	Bayer HealthCare
SKUP/2009/74	Glucose[1]	Accu-Chec Mobile	Roche Diagnostics
SKUP/2010/73	Leukocytes	HemoCue WBC	HemoCue AB
SKUP/2008/72	Glucose[1]	*Confidential*	
SKUP/2009/71	Glucose[1]	GlucoMen LX	A. Menarini Diagnostics
SKUP/2011/70*	CRP	smartCRP system	Eurolyser Diagnostica GmbH
SKUP/2008/69*	Strep A	Diaquick Strep A test	Dialab GmbH
SKUP/2010/67	Allergens	*Confidential*	
SKUP/2008/66	Glucose[1]	DANA DiabeCare IISG	SOOIL Developement co. Ltd
SKUP/2008/65	HbA1c	Afinion HbA1c	Axis-Shield PoC AS
SKUP/2007/64	Glucose[1]	FreeStyle Lite	Abbott Laboratories
SKUP/2007/63	Glucose[1]	*Confidential*	
SKUP/2007/62*	Strep A	QuikRead	Orion Diagnostica Oy
SKUP/2008/61	CRP	i-CHROMA	BodiTech Med. Inc.
SKUP/2007/60	Glucose[1]	*Confidential*	

*A report code followed by an asterisk indicates that the evaluation is not complete according to SKUP guidelines, since the part performed by the intended users was not included in the protocol, or the evaluation is a follow-up of a previous evaluation, or the evaluation is a special request from the supplier.

[1] Including a user-evaluation among diabetes patients