Comparative study of accuracy and clinical agreement of the CoaguChek XS portable device versus standard laboratory practice in unexperienced patients

Eduardo G. Torreiro; Elizabeth Gómez Fernández; Rosa Mariño Rodríguez; Carmen Vázquez López; Julia Barreal Núñez

Haematology Department, Monforte de Lemos Hospital, Monforte de Lemos, Spain

Summary
The objective of the study was to compare the accuracy and clinical agreement of the CoaguChek XS versus the standard laboratory practice. Forty-one patients on long-term anticoagulation with acenocumarol without previous experience in self-monitoring participated to obtain 218 pairs of data. Several methods for comparative statistics were applied to assess the possible disagreements between techniques as well as a range of previously published criteria of clinical agreement and the very recently described error-grid for INR comparison that we partially modify. The mean age was 52.1 and the indications for oral anticoagulation were prosthetic valves (36.59%), atrial fibrillation (34.15%), venous thromboembolic disease (21.95%) and others (7.31%) with a target range of 2–3 INR units (63.4%) or 2.5–3.5 (36.6%). Analyzing the whole series of data, the Pearson’s correlation coefficient for precision between methods was 0.95 and the C bias correction factor for accuracy 0.99 with a minimal bias of 0.1 INR units between methods applying the Bland-Altman plot. The linear regression procedure described by Passing and Bablok showed a minimal deviation from the best-fit line and a slope of 0.90. The mean of the absolute relative differences was 7% which is in the “very good” range of agreement. No results were found in the clinically “dangerous” D zone of the error-grids with 99% of data in the clinically irrelevant and low relevant areas A and B. In this study self-management with the CoaguChek XS was clinically safe and reliable.

Keywords
Self-management, oral anticoagulation, CoaguChek XS, acenocumarol, INR

Introduction
Because of the strong association between international normalised ratio (INR) levels and clinical outcome, maximal time spent on therapeutic range is crucial for oral anticoagulant therapy (1–4).

Achieving high-quality anticoagulation control can be difficult. The oral anticoagulant therapy monitoring requires frequent venous blood sampling and frequent trips to laboratories. Self-management of oral anticoagulant therapy facilitates the achievement of this goal.

Several studies have demonstrated that patients self-management of oral anticoagulant therapy can improve treatment quality. Surprisingly, in terms of clinical outcome, patient self-management achieved better results than conventional control, reducing the risk for major complications and exhibiting a trend towards reduced mortality. The self-monitoring yields a quick result, is readily performed and allows a proper dosing to get better clinical results and probably an overall economic saving (5–10) and, within an appropriate training program can improve patient satisfaction, disease knowledge and quality of life (11). The benefits of patient self-management in terms of INR measurements and time expended within range are also observed in patients aged 60 years or more (12).

The CoaguChek XS was released in 2005 to substitute the previous CoaguChek S system which had been validated in a number of clinical studies comparing its accuracy and clinical utility versus the standard laboratory practice (5, 13, 14) and although some studies have been conducted to demonstrate reproducibility of the results of the new CoaguChek XS versus the S series (15) there are still not too many publications comparing this device with the standard laboratory reference (16–20).
The objective of this study was to assess the accuracy of the INR determination performed by patients using the Coaguchek XS under real-life conditions by comparing the INR results of this monitoring device with those obtained at our standard reference laboratory for outpatients on oral anticoagulation and various pathologies (atrial fibrillation, mechanical heart valves, thromboembolic disease and others) as well to analyse the possible clinical repercussion of discrepancies.

Assuming that the veracity of the standard procedure is well documented we have applied comparative statistical techniques to analyse the degree of accuracy between the two procedures and clinical agreement techniques to study the clinical relevance of the possible differences (13, 22).

We place particular emphasis on analysis by the error-grid, recently adapted by Hemkens et al. (13) to be used in comparing clinical relevance of INR discrepancies. In that model we introduce a slight modification to define the high risk area, since our sample included patients in whom under-anticoagulation can be considered as dangerous as over-anticoagulation: mechanical prosthetic heart valves, chronic atrial fibrillation and cerebral ischaemia, etc (1, 2).

### Methods

**Patients and inclusion protocol**

This evaluation was undertaken as a part of an anticoagulation self-management pilot study and was conducted at the Monforte de Lemos Hospital, Spain, during the months from November 2007 to March 2008. Forty-seven ambulatory patients were initially included providing that their ages were between 18 and 70 and had an indication for long-term anticoagulation. The oral anticoagulant was acenocumarol (Sintrom®). Patients with lupus anticoagulant or relevant physical or psychological difficulties were excluded.

All the patients granted informed consent to participation and were assigned to a group of five to six people to attend a structured self-management program for patients on oral anticoagulation. The course consisted of a three hours session divided in two parts. The first lesson was taught by an experienced haematologist on anticoagulant therapy and included basic theoretical concepts about indications of long-term anticoagulation, INR concept and interpretation, factors that can influence the level of anticoagulation as well as identification and treatment of possible adverse events. The second lesson was conducted by a specially trained nurse about practical aspects for the correct collection of the sample, the use of the portable coagulometer and interpretation of the algorithms for dose adjusting. A great deal of practicing by the participants was undertaken at this moment. All these steps were conducted following the criteria of the previous ACOA (Alternative Control Oral Anticoagulation) clinical trial (5).

During the following weeks the patients came to the haematology laboratory facilities every three to seven days depending on the results of the previous test and the recommendations of the self-management algorithm. In every visit a venous sample was drawn to perform an INR analysis by our standard method and immediately afterwards the patients obtained the INR value by their own using the portable device. With that last value, the patients proposed a therapeutic scheme based on the algorithms. Once the patients were considered suitable for self-management with the criteria set by the haematologist and nurse on charge they entered the definitive self-management program at their homes. For the purposes of this study we used the paired data obtained during the weeks of the initial phase of the protocol.

**INR monitoring devices**

The Coaguchek XS is a portable coagulometer that measures the INR using whole blood obtained by finger prick and amperometric detection of the thrombin effect in the blood sample. It uses a human recombinant thromboplastin with an International Sensitivity Index (ISI) of 1.01 and every strip has an integrated quality control. Previous studies were performed to assign the ISI to the new test according to the World Health Organisation (WHO) guidelines for thromboplastins and plasmas used to control anticoagulant therapy, and to establish the measuring range of the new system (23–25). In our study, every patient used their own device with different lots of reactive strips as in real life conditions.

**Laboratory procedure**

The laboratory method consisted of phlebotomy drawing venous samples of whole blood into 0.109 M sodium citrate tubes at the laboratory facilities. Centrifuged samples were then analysed with a Sysmex CA1500 (Kobe, Japan). The thromboplastin used was Thromborel S (Dade International, Miami, FL, USA) from an only lot with an ISI of 1.04. All the samples were processed with the ordinary work of the laboratory in standard conditions.

**Statistical analysis**

Statistical analysis was performed using the computer program Medcalc v9.2. The agreement of the continuous data of the venous plasma and capillary prothrombin time measurement in INR units were examined using several methods for comparative statistics:

- We calculated the mean of the absolute difference \((|INR_{\text{reference}} – INR_{\text{self-monitoring}}|)\) and the relative difference \((|INR_{\text{reference}} – INR_{\text{self-monitoring}}| / INR_{\text{reference}})\) for the whole group of data, mean

<table>
<thead>
<tr>
<th>Table 1: Clinical characteristics of participants.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
</tr>
<tr>
<td>Mean age, years ± SD, (range)</td>
</tr>
<tr>
<td>Women n (%)</td>
</tr>
<tr>
<td>Time of previous anticoagulation, years ± SD (range)</td>
</tr>
<tr>
<td>Diagnosis, n (%)</td>
</tr>
</tbody>
</table>
- Mechanical prosthetic valve | 15 (36.59%) |
- Chronic atrial fibrillation | 14 (34.15%) |
- Venous thromboembolic disease | 9 (21.95%) |
- Others | 3 (7.31%) |
| Target range INR, n (%) | 15 (36.6%) |
- 2–3 | 26 (63.4%) |
- 2.5–3.5 | 15 (36.6%) |

INR units were examined using several methods for comparative statistics:

- We calculated the mean of the absolute difference \((|INR_{\text{reference}} – INR_{\text{self-monitoring}}|)\) and the relative difference \((|INR_{\text{reference}} – INR_{\text{self-monitoring}}| / INR_{\text{reference}})\) for the whole group of data, mean
absolute difference (MAD) and mean relative difference (MRD%), respectively.

The Bland and Altman plot (26) to calculate the difference between the results of the two methods plotted against their average, representing the limits of agreement between the two techniques the difference of the averages ± 1.96 standard deviation (SD).

The linear regression procedure described by Passing and Bablok (27) which gives an assessment of agreement of methods (slopes and coefficient of correlation).

Finally, the concordance correlation coefficient (28) to evaluate the degree to which pairs of observations fall on the 45º line through the origin and contains a measurement of precision $\rho$ and accuracy $C_b$, where $\rho$ is the Pearson correlation coefficient which measures how far each observation deviates from the best-fit line and is a measure of precision and $C_b$ is a bias correction factor that measures how far the best-fit line deviates from the 45º line through the origin and is a measure of accuracy.

**Clinical agreement of techniques**

We used three different published methods to assess clinical agreement in order to detect discrepant INR values that could have resulted in different clinical decisions.

Expanded agreement as described by Douketis et al. (21) was achieved if both the CoaguChek XS and the laboratory INR data were within, above or below the target range or the difference between both methods was no more than 0.5 units when only one of the pair of values was within the targeted range. This agreement is considered narrow if the first and fourth conditions are matched or in the case of values above the targeted range the difference is within 0.8 units and for values below targeted range the difference is no more than 0.5 units.

The accuracy rating according to Hill (22) applies to the MRD% being very good (6.58–9.25), good (9.32–11.86), acceptable (11.93–14.54), marginal (14.60–20.28) and very poor (20.34–26.99).

Finally we readapted the error grid approach initially described to evaluate the clinical significance of inaccuracies in self-monitoring of blood glucose in diabetes mellitus therapy (29–31) and adapted to compare INR measurements by Hemkens et al. (13). These authors define four risk zones that we partially modify and design a new error grid for patients with target INR between 2.5 and 3.5 units. The risk zones are defined as follows:

- **Zone A:** The clinical decision based on the pair of data is the same hence the difference between them is clinically irrelevant.
- **Zone B:** The corrective action according to the self-testing method would be in the correct direction but inadequate, is in fact unnecessary or a slight correction is needed although the patient seems to be in range by self-testing. These situations are considered of low clinical relevance since the corrective action would be rather small and the time to the next control will never be more than a week.
- **Zone C:** A substantial treatment correction is necessary according to self-testing although no modification would have been necessary by the reference method or the opposite situation. This is considered of moderate clinical relevance.
- **Zone D:** Dangerous situations in which the patient is exposed to a significant haemorrhagic or thromboembolic risk. The self-

| Table 2: INR results for the whole study group with the two methods. CoaguChek XS versus standard laboratory reference (MAD = mean of the absolute differences, MRD% = mean of the absolute relative differences). |
|---------------------------------|----------------|----------------|
| **INR mean ± SD**              | CoaguChek XS  | Standard Laboratory |
| 2.96 ± 0.95                    |                | 2.86 ± 0.83       |
| **INR range**                  | 1.30 – 6.5     | 1.26 – 5.37       |
| **MAD ± SD**                   | 0.21 ± 0.23    |                |
| **MRD (%) ± SD**               | 7 % ± 5.95 %   |                |

Figure 1: Bland-Altman plot diagram showing differences between the INR results for the CoaguChek XS and the standard laboratory method ($n = 218$, mean difference = 0.1, SD = 0.291).
testing and reference data differ substantially and the corrective action would be opposite to the necessary treatment, a corrective action is urgently needed but the self-measurement suggest being in the target range or the patient is in target range by the reference method whilst the self-management data drive to unnecessary important modification of therapy.

**Results**

After the initially training sessions six patients were excluded for not being considered suitable for self-management due to their own desire, difficulty to make a proper self-monitoring or to understand the self-management algorithms. The clinical characteristics of the remaining 41 participants are shown in Table 1.

In total, we obtained 218 paired data corresponding to these 41 patients with a number of determinations ranging between two and 15 and a median and SD of 5 ± 2.27. One hundred thirty-four of these results were from patients whose therapeutic range was 2–3 INR units and the other 84 from patients with target 2.5–3.5.

The mean INR, range, MAD and MRD% between methods are represented in Table 2.

The Bland and Altman plot (Fig. 1) shows a graphic representation of the difference (bias) when comparing the average of the two measurements: CoaguChek XS system and the laboratory reference. It shows a mean difference of 0.1 INR units and a mean relative bias of 2.7% being higher the values obtained with the CoaguChek XS. A visual inspection discloses a trend to greater differences in the INR range above 3.5 and especially over 4.5.

Applying the Passing and Bablok linear regression (Fig. 2) we observe a very slight deviation from the best-fit line with a slope of 0.90. This data is confirmed by a concordance correlation coefficient of 0.94 with a Pearsons precision \( \rho = 0.95 \) and a \( C_b \) bias correction factor for accuracy of 0.99.

<table>
<thead>
<tr>
<th>Error-grid risk zone</th>
<th>2–3</th>
<th>2.5–3.5</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>84.33%</td>
<td>79.76%</td>
<td>82.57%</td>
</tr>
<tr>
<td>B</td>
<td>14.93%</td>
<td>19.05%</td>
<td>16.51%</td>
</tr>
<tr>
<td>C</td>
<td>0.74%</td>
<td>1.19%</td>
<td>0.92%</td>
</tr>
<tr>
<td>Number</td>
<td>134</td>
<td>84</td>
<td>218</td>
</tr>
</tbody>
</table>

**Table 3: Percentage of pairs of data falling in every risk zone according to INR therapeutic range.**

The Bland and Altman plot (Fig. 1) shows a graphic representation of the difference (bias) when comparing the average of the two measurements: CoaguChek XS system and the laboratory reference. It shows a mean difference of 0.1 INR units and a mean relative bias of 2.7% being higher the values obtained with the CoaguChek XS. A visual inspection discloses a trend to greater differences in the INR range above 3.5 and especially over 4.5.

Applying the Passing and Bablok linear regression (Fig. 2) we observe a very slight deviation from the best-fit line with a slope of 0.90. This data is confirmed by a concordance correlation coefficient of 0.94 with a Pearsons precision \( \rho = 0.95 \) and a \( C_b \) bias correction factor for accuracy of 0.99.

**Figure 3: Error-grid for patients with INR target 2–3 and 2.5–3.5.**
When we proceed to analyse the different clinical agreement criteria described by Douketis et al. (21), 211 out of 218 pairs of data (96.79%) fitted into the narrow agreement and 214 of 218 (98.17%) fulfilled the expanded criteria of agreement, with only four pairs of data with no agreement according to these criteria.

In Table 3 we can observe the percentage of patients falling into every error-grid risk zone as defined above in the clinical agreement methods (Fig. 3). Notably, there were no data fitting risk zone D and a very high percentage of pairs (99.08%) fell into the A and B risk zones categorised as clinically irrelevant or low relevant differences. A very slight trend was observed for pairs of data fitting group B for patients in the higher INR therapeutic range.

**Discussion**

This comparative study of accuracy and clinical agreement of the CoaguChek XS versus our laboratory standard was part of a self-management pilot study in ambulatory patients. It is noteworthy that the only exclusion criteria were lupus anticoagulant or important physical or psychical disabilities unlike other studies in which there was a previous selection of patients to exclude those with important differences between the CoaguChek XS and the reference method (18). We included patients with very different clinical conditions and the two most frequent therapeutic ranges. Even more, the study was conducted in real-life conditions in anticoagulant control clinics, processing the samples as ordinary work and using different lots of reactive strips.

We used the Bland-Altman plot to assess the agreement between methods to demonstrate trends and systematic errors. With this comparative graphic we observed in our series of pairs of data a bias of 0.1 INR units between CoaguChek XS and our standard laboratory practice which is well within the range of the recommended analytical bias of ±0.20 INR units (32). The visual inspection of the diagram discloses an increase in the disagreement towards the higher INR averages, especially above 4.5. This is a foreseeable finding very well described in the previous bibliography studying the accuracy of portable devices (33, 34) and even between standard methods with different thromboplastins (35) and it is probably related to the WHO guidelines to exclude samples with INR outside the 1.5–4.5 range for ISI assignment (24) which makes the ISI/INR system accurate only in that range explored by the calibration procedure.

The MAD between pairs of data was of 0.21 INR units whereas the MRD was 7%. These figures are slightly better than the ones described for other portable devices (13) and fall into the “very good” range of agreement between methods (22).

The precision and accuracy between methods are finally confirmed by the results of the concordance correlation coefficient and the minimal deviation from the best-fit line in the Passing-Bablok linear regression method.

In order to assess the possible clinical consequences of disagreements we applied the different criteria previously described. The vast majority of the pairs of data (98.17%) fulfilled the Anderson accuracy criteria (36) or the expanded criteria of agreement, whereas 96.79% fitted the narrow agreement of Douketis et al. (21). These results are slightly worse than the 100% agreement for another comparison between the same methods (16), what could probably be related to the fact that we did not previously exclude patients with important differences and we also had an important amount of high INR data whilst in that study there is only one value over 3.5.

The only four pairs of data not fitting the accuracy criteria corresponded to INR values in the upper limit of the therapeutic range as determined by the reference method and a comparative overestimation of 0.52 to 0.81 INR units by the CoaguChek XS which was more than the 0.4 or 0.5 INR units required by the Anderson or Douketis criteria in this scenery. This situation led to a very slight diminution of the anticoagulation dose without any clinical consequences since in every case the patients were within the therapeutic range with both methods at the next monitoring one week later.

We designed a new error grid for the higher 2.5–3.5 INR range and made an important modification of the grid described by Hemkens et al. (13) transforming their C3-risk zone into D-high risk-zone. This modification corresponds to situations where the overestimation of the INR by the portable device could lead to therapeutic attitudes resulting in under-anticoagulation and consequently increasing the thromboembolic risk. This is particularly significant in our study population with a high percentage of patients with mechanical prosthetic heart valves, secondary prophylaxis for chronic atrial fibrillation and previous embolic incidents or high risk venous thromboembolic disease in patients who suffered several events and/or complex genetic hypercoagulability predisposing mutations.

No results were found in the clinically “dangerous” zone D, and additionally only about 1% were detected in the clinically moderate relevant zone C with 99% of data in the clinically irrelevant and low

**What is known about this topic?**
- Self-management of oral anticoagulation in the context of scheduled training programmes achieves a reduction in major haemorrhagic and thromboembolic complications when compared with conventional control.
- There are minimal deviations between international normalised ratio (INR) measurement between self-monitoring coagulometers and laboratory methods, especially in the higher INR ranges.

**What does this paper add?**
- The accuracy and agreement between CoaguChek XS and standard laboratory practice was very good in a population of patients with very different clinical conditions and no previous experience in handling self-monitoring devices.
- The error-grid complements the statistical calculations and contributes to the study of variations in therapeutic interventions and clinical relevance of results differences between INR measurement devices and methods.
- The CoaguChek XS was highly accurate, with minimal differences with the reference laboratory method in the higher INR levels as expected, but no clinical relevant consequences.
relevant zones A and B. Very interestingly, the four pairs of data not fulfilling the previously described agreement criteria were within the B risk zone of the error grid corresponding to situations considered of low clinical relevance. We consider that the error-grid approach represents an improvement for the evaluation of clinical agreement between methods, especially for those cases when only one data of the pair is within the targeted range. Moreover it establishes a continuum of risk stratification instead the yes or no of previous approaches and should be validated in clinical studies to assess the possible correlation of the different risk areas with clinical outcomes. This would also help to line out the boundaries of the risk zones using clinical criteria.

In summary, the CoaguChek XS was highly accurate, with minimal differences with the reference laboratory method in the higher INR levels as expected, but no clinical relevant consequences. This is especially noteworthy if we take into account the varied population studied with different INR therapeutic ranges and no previous experience in handling self-monitoring devices. We conclude that patient self-management using CoaguChek XS in this study was clinically safe and reliable and for clinical purposes INRs can be considered interchangeable between the laboratory and the CoaguChek XS.

References