Safety and Effectiveness of Point-of-Care Monitoring Devices in Patients on Oral Anticoagulant Therapy: A Meta-analysis

Authors:

*†‡ Philip S. Wells MD MSc, **Allan Brown MA MBA, †‡James Jaffey MA MSc, **Lynda McGahan MSc, § Man-Chiu Poon MD MSc, **Karen Cimon MLT

Institutional Affiliations:

1.1 *Department of Medicine, University of Ottawa.

1.2 †Department of Epidemiology and Community Medicine, University of Ottawa.

1.3 ‡Ottawa Health Research Institute.

1.4 **Canadian Agency for Drugs and Technologies in Health, Ottawa, Canada

§ Department of Medicine, University of Calgary and Foothills Medical Centre

Corresponding Author:

Dr. Philip Wells

Chief, Division of Hematology

Canada Research Chair in Thromboembolic Disease

Ottawa Hospital, Civic Campus

Suite F649; 1053 Carling Avenue

Ottawa, Ontario; K1Y 4E9, Canada

Tel.: (613) 798-5555 ext. 18769; Fax: (613) 761-5351

E-mail: pwells@ohri.ca

Grant Support:

Dr Wells is the recipient of a Canada Research Chair Award.

Word count of abstract: 274
Word count of text: 3869

Background: Point-of-care devices (POCDs) for monitoring long-term oral anticoagulation therapy (OAT) may be a useful alternative.
Purpose: To determine clinical outcomes using POCDs for OAT management by performing a meta-analysis. Previous meta-analyses on POCDs have serious limitations.
Data Sources: PubMed, The Cochrane Library, DIALOG®'s MEDLINE®, EMBASE®, BIOSIS Previews® and PASCAL databases
Study Selection: Randomized controlled trials of patients on long term OAT; anticoagulation monitoring by POCD compared to laboratory INR and clinic management.
Data Extraction: 1) rates of major hemorrhage; 2) major thromboembolic events rates; 3) percentage of time within therapeutic range;4) deaths. Outcomes were compared using a random-effects model. Summary measures of rates were determined. Quality of studies was assessed using the Jadad Scale.
Data Synthesis: Sixteen articles (15 studies) were included. POCD INR testing reduced the risk of major thromboembolic events (OR= 0.48; 95% CI 0.33 to 0.72) and deaths (OR=0.54; 95% CI 0.35 to 0.83), and resulted in better INR control compared to laboratory INR testing. The odds ratio for major hemorrhage was not significantly different.
Limitations: Quality scores varied from one to three (out of a maximum of five).Only three studies defined how thromboembolic events would be diagnosed casting doubt on the accuracy of thromboembolic events. The studies suggest only 24% of patients are appropriate for self-testing/self-management. POCD patients underwent INR testing at a much higher frequency and received much more intensive education on OAT management.
Conclusions: Use of POCDs is safe and may be more effective. However, most patients are not appropriate for self-testing/self-management. Patient education and frequency of testing may be the most important factors. Definitive conclusions about the clinical benefits provided by self-testing/self-management require more rigorously designed trials.

Oral anticoagulants in the form of vitamin K antagonists, are widely used for the prevention and treatment of thromboembolic events in various clinical conditions. Long-term use is typically required for high risk groups with particular conditions such as mechanical heart valves, chronic atrial fibrillation, venous thromboembolism, acute myocardial infarction, stroke, and peripheral arterial occlusion.(1,2) For many of these indications, a person must continue on oral anticoagulant therapy (OAT) lifelong.(3,4) Because of the aging population, and an associated increase in the prevalence of atrial fibrillation and venous thromboembolism, it is expected that more patients will need OAT in the future. Evidence suggests that OAT reduces the incidence of thromboembolic complications (venous and arterial thrombosis), and associated mortality and morbidity in these patient populations. (5) Vitamin K antagonists have a narrow therapeutic window. Excessive anticoagulation confers an increased risk of bleeding, while sub-therapeutic anticoagulation is associated with an increased risk of stroke and other thromboembolic events. (6,7) Unfortunately, the biological effect of the vitamin K antagonists varies between individuals and within an individual over time.(5) For this reason, patients need regular monitoring of OAT with a test known as the international normalized ratio (INR), usually determined in a hospital or outpatient laboratory facility by a venipuncture sample processed in the lab. This can be inconvenient in terms of the blood sampling procedure and of the time spent going for a laboratory test.(1,2) Point-of-care devices (POCDs) for monitoring long-term OAT were introduced in the 1990s. POCDs are portable and only require a drop of blood from a fingertip puncture. In some countries, such as Germany, self-testing and self-management with POCDs are widely employed, but in most countries uptake has been limited.(8,9)

The POCD technology makes it possible for patients on long-term OAT to self monitor and self manage their OAT. Those who manage OAT programs need to know how POCDs compare in effectiveness and cost effectiveness with standard laboratory tests. The objective of this study was to perform a meta-analysis to assess the clinical implications of POCD use for OAT monitoring and also to assess any potential limitations with the data.

2 Methods

3 Literature search strategy

We obtained published literature by cross-searching DIALOG®'s MEDLINE®, EMBASE®, BIOSIS Previews®, and PASCAL databases. There were no year or language restrictions. A broad search strategy with appropriate descriptors and keywords was used, in combination with a filter, to restrict results to controlled trials, meta-analyses, and systematic reviews. We also ran parallel searches on PubMed and the Cochrane Library.

The original search was performed in July 2005. Regular alerts were established on MEDLINE®, BIOSIS Previews®, and EMBASE® databases to capture new studies up to August 11, 2006 and searches in the Cochrane Library were updated regularly. We obtained grey literature by searching the web sites of regulatory agencies, health technology assessment agencies, and near-technology assessment agencies. Specialized databases, such as the University of York's NHS Centre for Reviews and Dissemination and the Latin American and Caribbean Center on Health Sciences Information (LILACS), were also searched. The following professional associations' web and conference sites were searched for additional information: Thrombosis Interest Group of Canada, Canadian Cardiovascular Society, American College of Cardiology, American Society of Hematology, and European Society of Cardiology. Non-randomized controlled trials were included in the literature search because of their potential use in other sections of the report.

Selection criteria and method

Studies that were included met the following selection criteria:

- study design: randomized controlled trial (RCT)
- population group: patients on long-term (at least three months) OAT (no a priori restrictions on age or mental capacity)
- interventions: anticoagulation monitoring by POCD; this could include POCD testing at an anticoagulation clinic, POCD self-testing by the patient, POCD self-testing plus self-management and control, or any other POCD management strategy
- comparators: usual care (venipuncture blood draw for an INR lab test and management provided by an anticoagulation clinic or individual practitioner)
- outcomes: for studies to be included, they were required to have reported on at least one of the following:
    1. Rates of major hemorrhage, where major was defined as resulting in death, or hemorrhage was clinically overt and showed one of the following: critical site involvement (intra-cranial, retroperitoneal, intra-ocular, intra-spinal, or pericardial), drop in hemoglobin of ≥2.0 grams per deciliter, need for transfusion of >2 units of packed red blood cells, or a bleeding index of >2.0.
    2. Major thromboembolic event rates, noting whether the study required objective diagnostic tests for venous and arterial thromboembolic complications. Transient ischemic attacks were considered to be minor thromboembolic events and were included in a secondary analysis to evaluate all thromboembolic events.
    3. Percentage of time the patient's blood was within the normal therapeutic INR range according to a method described by Rosendaal *et al.*(10) The Rosendaal algorithm is used to calculate the time that a patient stayed in a pre-determined INR interval. The algorithm assumes a linear increase or decrease between two consecutive INR determinations.(10)

Reports were excluded if they were duplicate reports, preliminary reports of data presented in full, dose-finding studies, studies in which oral anticoagulants were combined with antiplatelet drugs, and studies that did not follow patients for more than three months. While we had planned to exclude data based on patients who had not been on OAT for three months upon entering the study, we dropped this criterion and performed analyses with and without these studies.

We assessed the retrieved references for possible inclusion based on an evaluation of the title and the abstract according to the selection criteria. The reviewers pilot-tested the inclusion-exclusion criteria on seven articles and performed a calibration exercise to ensure consistent application. Letters to the editor, review articles, editorials, and commentaries were excluded. The remaining studies were fully assessed.

At least two reviewers independently reviewed each citation from the literature search. At the first stage, abstracts were selected independently by KC and LM. Consensus was reached by discussion. At the second stage, full-text articles were reviewed independently. Agreement on eligibility was achieved by discussion between the two reviewers.

Data extraction

A data extraction form was developed a priori. Two reviewers (PW, KC) independently extracted data from eligible articles and assessed their quality using a standard electronic form. PW and

another reviewer (LM) then arrived at a consensus on the extracted data and quality values through discussion.

Strategy for quality assessment

Study quality was assessed using the criteria proposed by Jadad *et al*., and the adequacy of allocation concealment was evaluated as appropriate or inappropriate according to the criteria proposed by Schulz and Grimes.(11,12) If information in the reports was insufficient, these issues were recorded as unclear or unstated. We successfully contacted authors when data were incomplete or missing.

Data analysis methods

For assessing the outcomes of major hemorrhage, major thromboembolic events, and all thromboembolic events, we conducted a meta-analysis by calculating odds ratios and their 95% confidence intervals (CI) for the event rates, comparing results between POCD testing and laboratory testing. A comparison of death rates was also performed. We used a random-effects model for all comparisons (according to the method described by DerSimonian and Laird (13)), recognizing that its use can reduce the effect of larger studies relative to a fixed-effects model. A random-effects model allows for between-study variation, and was chosen as the more conservative option. Differences between effects were tested using a Z test, and p values <0.05 were considered to be significant. Our plan was to analyze the differences in patient characteristics and other study differences, to evaluate the cause of any heterogeneity. For each comparison group, we estimated the between-study heterogeneity using the Q statistic in the Review Manager (RevMan) software package. Heterogeneity was considered significant for $p<0.05$. The I^2 statistic, indicating the proportion of total variation due to heterogeneity, was also calculated. For I^2, the cut-points were 25%, 50%, and 75% for low, moderate, and high heterogeneity, respectively.(14) The summary measures of rates of major hemorrhage and thromboembolic events were determined using the inverse variance weighted averages. Forest plots were prepared. Funnel plots were generated to assess whether the magnitude of the observed association was related to the variance of each study and whether there was evidence of publication bias. We did a paired t-test of mean percentage time in range for the control and intervention groups.

Our analysis was performed for five groups:

- all eligible studies
  - studies that required patients to be on OAT for >3 months before study entry, analyzed as a subgroup: these patients are more likely to stay on OAT, because they are familiar with what is involved in OAT care; any events in the first three months are less likely to be related to OAT control
  - studies that compared POCD self-testing and self-management to routine care: these look at patients' use of POCDs rather than use of POCDs by health care professionals
  - studies that described the requirement for objective tests to diagnose major thromboembolic events: these are most likely to have the most accurate data
  - studies scoring ≥3 on the Jadad assessment tool: these use better methods, and we can evaluate whether poorly designed studies are driving the results produced when all studies are included.

Results

Quantity of research available

We identified 439 citations in our initial search. Routine updates yielded an additional 13 citations for a total of 452 (Figure 1). Of these 452 citations, 409 did not meet the selection criteria and were

excluded, leaving 43 (39 from the initial search and four from the updated searches). Two were added for reconsideration after the study criteria were revised to include studies where patients had been on OAT for <3 months at the start of the study. We retrieved 45 potentially relevant articles for further review. Of these, we excluded 29 articles(15-43), leaving 16 relevant articles describing 15 unique RCTs.(44-59) One RCT was reported in two publications.(50,51) These articles by Koertke *et al.* were not duplicates because they reported different aspects of the RCT but informed the data extraction for the same study. The Gadisseur article provided two sets of data because the authors compared self-test plus self-management and self-test plus clinic management to routine care.(47)

Of the 11 articles that were excluded based on study design, four were reviews.(22-25) Five articles were excluded because the intervention used was inappropriate for our review.(26-29,43) For example, while POCD testing was used in some studies, the patients were managed based on results from laboratory testing, not POCD testing. As a result, no true comparison could be made with those patients in a group undergoing laboratory testing because this study design may miss results related to management based on POCD testing. One article was excluded based on population (30), one based on outcome measures (36), and three because they were at the protocol stage.(31-33) Three were duplicates of excluded articles (34,35,37), and five were duplicates of included articles, so they were also excluded.(38-42)

Study characteristics

Table 1 summarizes the characteristics of the studies and demonstrates their variability regarding observation periods, mean age of patients, and indication for anticoagulation. It was not possible to break down study outcomes according to the indication for anticoagulation. We found 11 studies that compared self-monitoring plus self-management to routine anticoagulation control. In eight studies, only patients who had been on OAT for ≥3 months were enrolled; in seven studies, patients were enrolled from the time of initiation of anticoagulation or the time could not be determined. (46,50,51,53,56-59) Table 2 summarizes the outcome data from eligible trials.

Data analysis and synthesis

The intervention group comprised 2,144.6 patient-years of observation, while the control group comprised 2,316.1 patient-years. For all studies, there were significantly fewer major thromboembolic events in the POCD testing group than in the routine care group (OR=0.48; 95% CI 0.33 to 0.72). This statistically significant difference was also observed in all four of the other subgroups (Table 3). The odds ratios for all thromboembolic events were similar. Death from any cause was significantly less likely in the POCD testing group (OR=0.54; 95% CI 0.35 to 0.83) when all eligible studies were pooled, and this remained significant in all other analyses except the ones that included only three studies (i.e., those that defined the objective diagnostic criteria). For major hemorrhage, the odds ratio was not significantly different between the POCD testing group and the routine testing group in any of the analyses (OR=0.75; 95% CI 0.51 to 1.10 for the analysis of all studies).

The percentage of time in range was significantly better for the POCD group in all four relevant analyses (Table 3). For the "all studies" analysis, the mean percentage of time in range for the POCD testing group was 69% (95% CI 65% to 72%) versus 61% (95% CI 55 % to 66%; p=0.004) for the routine care group. In the analysis comparing self-testing plus self-management with routine testing, the means were 71% (95% CI; 66% to 76%) versus 63% (95% CI; 60% to 66%), p=0.016, respectively.

Figures 2 to 5 show forest plots for major hemorrhage, major thromboembolic events, all thromboembolic events, and death. Figure 6 shows the funnel plot for all thromboembolic events, which appears to be symmetrical and does not give an indication of publication bias. Figure 7 shows a funnel plot for major hemorrhage that suggests the possibility of publication bias. The funnel plots assumed a fixed effect, because the software used does not allow random effects for funnel plots. Figures 2 to 5 also provide information for assessing heterogeneity using Q and I^2 statistics. This indicates a small effect of heterogeneity on the meta-analysis.

We had hoped to be able to compare quality of life (QoL) scores between studies, but QoL was not uniformly measured, and when it was, different tools were used. Five studies planned and performed formal evaluations. One used the EuroQol (49) and reported no significant changes or differences between the study groups from study inception to completion. Two used the same 40-item structured questionnaire.(44,53) Cromheeke demonstrated significant differences in five categories, in favour of the self-management group, and Sawicki demonstrated similar findings with the most pronounced improvements in general treatment satisfaction scores and distress scores. Two studies used locally developed satisfaction scales and demonstrated that patients were satisfied using POCDs, but the studies did not do any formal comparisons.(54,59)

The quality score of the 15 studies varied from one to three, with eight attaining a score of three out of a maximum of five (Table 2). Two studies received a score of one. Although no study was double-blinded (it could be argued that this is reasonable, so the maximum quality score would be three), the investigators could minimize potential bias by evaluating outcomes of hemorrhage and thromboembolic events without knowing whether patients underwent POCD testing or laboratory testing. This was done in four studies.(47,53,57,59) Five studies used adequate allocation concealment.(44,45,52,58,59) Only three studies stated and defined how thromboembolic events would be diagnosed.(52,54,58) This casts doubt on the accuracy of the recorded number of thromboembolic events. Most studies did define a priori the criteria for major hemorrhage.

To evaluate the potential scope for the use of POCDs, we considered whether they were well tolerated and easily employed. We looked for data on patient eligibility, agreement to consent, and withdrawal from studies. This is most relevant for the 11 studies in which patients were using the POCD for self-management. Three studies did not provide these data.(48,56,58) The studies that did report these data showed the following:

- For the proportion of patients deemed eligible from a group of consecutive patients, eight studies demonstrated 16% to 40% of patients were deemed unsuitable to use the POCD device, with most studies reporting closer to 40% as unsuitable.
- In six of the 11 studies, 18% to 28% of the patients dropped out after being randomized to use the POCD and attempting the training program.
- Six studies reported that 12% to 19% abandoned the use of the POCD after the study began, compared with 0% to 6% who withdrew in the routine care groups.

Discussion

Our study reviewed 15 RCTs comparing POCD testing to routine anticoagulation monitoring care in a hospital or laboratory. Of those, 11 compared using POCDs for self-management and self-monitoring with routine anticoagulation monitoring. The latter comparison is the most relevant from the perspective of patient's convenience of care, and the data in these trials suggest that POCDs resulted in significantly fewer major thromboembolic events and deaths from any cause. The risk of major hemorrhage was similar between the two groups. These results were unchanged for all

subgroup analyses performed. For all analyses, the comparison of percentage time in range between the POCD group and the control group demonstrated superiority with POCD use. Three studies described the use of objective tests to diagnose thromboembolic events.

To test the robustness of the data we performed several subgroup analyses. One included only patients who had been on OAT for ≥3 months. We were initially concerned that including patients who were not yet stable on OAT could bias the results against the use of POCDs, given the known interactions with heparin and the difficulty in first achieving INR control.(60) Our subgroup analysis showed that the results were essentially the same, regardless of the time that patients were on OAT at baseline. The odds ratios were also similar whether we included all studies (i.e., any POCD testing compared to routine INR testing) or just self-test plus self-management comparisons. The outcomes were unchanged regardless of who performed the dosing in the POCD groups.

The summary data suggest that POCDs are advantageous. However, in all studies, the frequency of INR monitoring was higher in the POCD group than with routine anticoagulation monitoring. In most cases, the frequency of testing was dictated by the study protocol. It remains unknown whether similar frequent monitoring in routine care would eliminate these differences. It is also unknown whether this rigorous frequency of monitoring using POCDs would persist outside the study setting. In addition, patients who self-manage lose regular contact with their physician. The implications of this are unknown, but based on the critical endpoint of death, it does not seem to be a disadvantage. Of note, is that we calculated odds ratios to approximate the relative risk, since the event rates were relatively "rare" (some take this as <10%) so the odds ratio approximation of the relative risk is good. The OR will always be further from the neutral point of 1 than the relative risk (i.e., less conservative), so the results should be interpreted with this in mind.

These findings are subject to certain limitations. First, study methods were not ideal. The highest quality score for any of the publications was three, no study was double-blinded, and in many studies it was impossible to categorize what happened to the withdrawals. In all studies, the withdrawal rates of the POCD testing groups were higher than those of the routine testing groups. In most studies, thromboembolic events were not evaluated in a blinded and objective manner. This introduces the risk that the summary estimates may be biased, but deaths were also statistically significantly less frequent in the POCD group and death is an objective measurement. Secondly, the criteria for patients' eligibility for inclusion into the individual clinical trials and the high withdrawal rates are such that it is difficult to determine generalizability. It seems that the inclusion-exclusion criteria for the individual trials resulted in the exclusion of many patients deemed unsuitable for POCD self-testing before randomization, in addition to many such patients declining the invitation to participate. Perhaps more importantly, many patients failed to complete POCD training and many who completed training subsequently dropped out. An upper bound of 24% of OAT patients could be eligible for self-testing or self-management with POCD.(36) Consequently, the results of our meta-analysis may only apply to selected patients. Third, the INR test frequency was much higher in the POCD testing strategies; INRs were performed approximately weekly versus monthly in the standard group. It remains unknown whether similar frequent monitoring in routine care would eliminate these differences. It is also unknown whether this rigorous frequency of monitoring using POCDs would persist outside the study setting. Furthermore, the POCD group underwent three two-hour, small group education sessions in most studies whereas no special education was provided to the standard care groups. In addition, patients who self-manage might lose regular contact with their physician. The implications of this are unknown, but based on the critical endpoint of death, it does not seem to

be a disadvantage. These limitations make it impossible to determine what effects the frequency of monitoring, patient education, the POCD, or the patient selection may have on the outcomes.(60)

While information on QoL and patients' satisfaction with the POCDs was collected, we were unable to provide quantitative summary measures of these factors. A qualitative analysis of the data suggests that, in general, patients were at least as satisfied, and some more so, with self-managing using the POCDs than with receiving care at an anticoagulation clinic. These patients were good candidates for using a POCD, having been self-selected or selected by a health care researcher.

This report is not the first systematic review to compare POCD testing with laboratory testing for the management of patients on OAT. Five systematic reviews have been published, but all have limitations.(22,24,61,62) The most recent analysis by Heneghan *et al*. suggests that POCD testing resulted in significantly fewer major hemorrhages, thromboembolic events, and deaths than conventional INR testing.(61) The Heneghan analysis, which comprised 14 studies, included three that we considered to be ineligible: one compared POCD testing only with other POCD testing (i.e., inappropriate comparison for their analysis), one did not use POCD test results for OAT management, and the third followed patients for only eight weeks.(21,27,43) Our report also included four articles that were excluded by Heneghan.(45,46,54,58) Finally, the summary data used by these authors reported events per patient enrolled and not by length of follow-up, as we did. As such, the rates we report will be more comparable. Despite this, we came to similar conclusions with respect to two of four outcomes; death and all thromboembolic events. The other authors did not report separately on the more relevant outcome of major thromboembolic events. We did not detect a significant difference between groups for the outcome of major hemorrhage.

Our meta-analysis suggests that using POCDs to manage OAT results in significantly fewer deaths and thromboembolic events and better INR control than laboratory INR testing. Under usual care, warfarin therapy requires regular laboratory monitoring of the INR, coupled with frequent physician-patient contact for dosage adjustment to ensure efficacy and safety.(9) The usual-care method can be cumbersome and inconvenient for the patient and the physician. There is also a potential for dosing errors due to misinterpretation of information conveyed by the physician or delays in contacting the patient.(9) This, plus faster test results, more convenience, and more frequent testing, are plausible reasons for the incremental health benefits observed with POCD use.

Unfortunately, the studies designed to date do not allow a determination of why POCD care is superior. To determine whether it is patient education or frequency of testing that provides superior outcomes requires further randomized controlled trials. Widespread adoption of POCD monitoring at this time would be premature.

References