

A comprehensive evaluation of rhythm monitoring strategies for the detection of atrial fibrillation recurrence: insights from 647 continuously monitored patients and implications for monitoring after therapeutic interventions

Article (Unspecified)

Charitos, Efstratios I, Stierle, Ulrich, Ziegler, Paul D, Baldewig, Malte, Robinson, Derek, Sievers, Hans-Hinrich and Hanke, Thorsten (2012) A comprehensive evaluation of rhythm monitoring strategies for the detection of atrial fibrillation recurrence: insights from 647 continuously monitored patients and implications for monitoring after therapeutic interventions. *Circulation*, 126 (7). pp. 806-814. ISSN 0009-7322

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/47108/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A Comprehensive Evaluation of Rhythm Monitoring Strategies for the Detection of Atrial Fibrillation Recurrence: Insights From 647 Continuously Monitored Patients and Implications for Monitoring After Therapeutic Interventions

Efstratios I. Charitos, MD; Ulrich Stierle, MD; Paul D. Ziegler, MS; Malte Baldewig, MD; Derek R. Robinson, MA, MSc, DPhil, CStat; Hans-Hinrich Sievers, MD; Thorsten Hanke, MD

Background—Intermittent rhythm monitoring (IRM) to detect atrial fibrillation (AF) recurrence is employed to evaluate the success of therapeutic interventions. In a large population of patients with continuous monitoring (CM), we investigated the sensitivity of various frequencies and durations of IRM strategies on the detection of AF recurrence, the dynamics behind AF recurrence detection, and we describe measures to evaluate temporal AF recurrence.

Methods and Results—Rhythm histories of 647 patients (mean AF burden, 0.12 ± 0.22 ; median, 0.014; 687 patient-years) with implantable CM devices were reconstructed and analyzed. With the use of computationally intensive simulation, the sensitivity of IRM of various frequencies and durations on the identification of AF recurrence was evaluated. Prolonged-duration IRM was superior to shorter IRM ($P < 0.0001$). However, even with aggressive IRM strategies, AF recurrence was not detected in a great proportion of patients. The temporal AF burden aggregation (AF density) was directly related to IRM sensitivity ($P < 0.0001$). Even at similar AF burdens, patients with high-density AF required higher-frequency or prolonged-duration IRM to achieve the same sensitivity as in low-density AF ($P < 0.0001$). Patients with high-density, low-burden AF benefit the most from CM for detection of AF recurrence.

Conclusions—IRM follow-up is significantly inferior to CM. IRM strategies will not identify AF recurrence in a great proportion of patients at risk. Temporal AF characteristics play a significant role in AF recurrence detection with the use of IRM. For the scientific, evidence-based evaluation of AF treatments, CM should be strongly recommended. Prospective studies are required to evaluate whether CM to guide clinical management can also improve patient outcomes.

Clinical Trial Registration—URL: <http://www.clinicaltrials.gov>. Unique identifier: NCT00806689. (*Circulation*. 2012;126:806-814.)

Key Words: atrial fibrillation ■ atrial fibrillation arrhythmia ■ rhythm monitoring ■ rhythm recorders

Detection of atrial fibrillation (AF) recurrence after therapeutic interventions has until now been based primarily on the results of intermittent (“snapshot”) rhythm monitoring (IRM). This approach has significant limitations in terms of sensitivity because the monitoring period of these examinations is limited; however, the results of these examinations are used to draw inferences on the success of ablation procedures or pharmacological strategies. It has been shown that reliance on symptoms or office ECGs overestimates the success rate of the ablation procedures and will misclassify patients who have recurrence of paroxysmal AF as being in sinus rhythm.^{1–4} The current consensus on AF monitoring recommends at least two 24-hour Holter monitor (HM)

examinations annually for the detection of AF recurrence after ablation procedures,^{5–7} which, however, has also been shown to underdetect AF recurrence and thus overestimate procedural success.^{1,8} Reliable and accurate detection of AF recurrence is thus of special importance for the evaluation of pharmacological or ablation therapies as well as when decisions on changes in anticoagulation or antiarrhythmic therapy are to be made.

Editorial see p 791 Clinical Perspective on p 814

With the introduction of implantable, leadless rhythm recorders, continuous monitoring (CM) has been proposed

Continuing medical education (CME) credit is available for this article. Go to <http://cme.ahajournals.org> to take the quiz.

Received February 8, 2012; accepted June 14, 2012.

From the Department of Cardiac and Thoracic Vascular Surgery, University of Luebeck, Luebeck, Germany (E.I.C., U.S., M.B., H.S., T.H.); Medtronic Inc, Minneapolis, MN, United States (P.D.Z.); and Department of Mathematics, School of Mathematical and Physical Sciences, University of Sussex, Brighton, United Kingdom (D.R.R.).

Correspondence to Efstratios I. Charitos, MD, Department of Cardiac and Thoracic Vascular Surgery, University of Luebeck, Ratzeburger Allee 160, 23568 Luebeck, Germany. E-mail efstratios.charitos@gmail.com

© 2012 American Heart Association, Inc.

Circulation is available at <http://circ.ahajournals.org>

DOI: 10.1161/CIRCULATIONAHA.112.098079

recently for more accurate detection of AF recurrence after ablation procedures or novel pharmacological strategies.⁹ As expected, the sensitivity of these diagnostic modalities is close to 100% and, as such, is much superior to that of the currently employed 24-hour HM strategies.^{8,10} However, the implantation of a CM device in all patients may seem unrealistic. This led to the belief that intensifying noninvasive IRM either with greater frequency or with longer duration (7- or 30-day monitoring) or both may lead to better and more reliable detection of AF recurrence.^{1,11,12}

Recently, data from CM patients have been used to correlate the recurrence of AF with the risk for thromboembolic complications.^{13–16} The advance of this technology has not only provided better insight into the rhythm evaluation of these patients but has also initiated a change of mentality regarding AF from a qualitative (“yes/no”) to a quantitative (amount of AF) approach. The AF burden, defined as the proportion of the total monitored time a patient is in AF, has been utilized in evaluating the risk for thromboembolism in patients with AF.^{13–15} However, even these measures present a rather static approach in the evaluation of the dynamic incidence of AF recurrence, partially ignoring the temporal dispersion or temporal aggregation of the AF episodes and AF burden.

The aim of the present study was 2-fold: (1) to identify the sensitivity of IRM strategies of various frequencies and durations on the accurate detection of AF recurrence with the use of data collected from patients with heart rhythm CM devices and (2) to investigate the dynamics behind AF recurrence detection and propose measures to evaluate the temporal pattern of AF recurrence.

Methods

Data acquired from 647 patients monitored with a CM device (Reveal XT 9529, n=73; AT500 pacemaker, n=574; Medtronic, Inc, Minneapolis, MN) were analyzed. Demographics and patient characteristics are summarized in Table 1. All patients provided informed consent for the data collection and use, and the study was approved by the local ethics committee.

With the use of data from the CM device, all AF episodes were examined. Isolated AF episodes of <5 minutes duration that could present artifacts were disregarded, and patients with no documented AF episode were regarded as AF free (n=174). The complete rhythm history of every patient was reconstructed (Figure 1). Thereafter, we calculated the probability of successful identification of AF recurrence by IRM of various durations (eg, 24 hours and 7, 14, and 30 days) and frequencies for every patient. Computationally intensive simulation was employed to simulate in every patient all possible IRM strategies of various durations and frequencies to draw inferences on the sensitivity of the monitoring strategy and burden characteristics of the patients being monitored. Sensitivity was defined as the proportion of patients correctly identified as having AF recurrence with the simulated IRM strategy to the true number of patients having AF recurrence (identified from CM). AF burden was defined as the proportion of the time in AF in relation to the total monitored time. The simulation trials were performed as follows: After reconstructing the rhythm history of every patient *j*, monitored for a total of *g* days, we defined the sample space $\Omega_{kj} = \{1, 2, \dots, g-k+1\}$ to be the set of possible days that a *k*-day intermittent monitoring could be started (for *k*=1, 2, 7, 14, and 30). A *k*-day monitoring starting on day *i* $\in \Omega_{kj}$ therefore included the following associated monitored days: *i*, *i*+1, ..., *i*+*k*-1. To simulate *n* independent *k*-day monitorings of patient *j*, *n* elements were selected at random from Ω_{kj} , except that elements were rejected if their

Table 1. Demographics of the Patient Population

	Total	%
Male	376	58.1
Age, mean±SD, y	68.9±12.3	
Follow-up, mean±SD, range, y	1.1±0.4, 0.1–3.7	
History of atrial arrhythmia		
Atrial tachycardia	114	17.6
Atrial flutter	176	27.2
Paroxysmal AF	475	73.4
Persistent AF	32	4.9
Long-lasting persistent AF	35	5.4
History of cardioversion	18	2.8
Cardiovascular history		
Ischemic heart disease	99	15.3
Coronary artery disease	220	34.0
Cardiomyopathy	64	9.9
Hypertension	405	62.6
History of ablation for AF		
Cox/Maze III procedure	17	2.6
Left sided only	53	8.2
AV node ablation	28	4.3
Other	71	11.0
History of cardiac surgery		
CABG	113	17.5
MVR	45	7.0
AVR	41	6.3
TVR	7	1.1
Ascending aorta replacement	9	1.4
PVR	1	0.2
NYHA class		
I	331	51.2
II	234	36.2
III	67	10.4
IV	3	0.5
Pacing indication		
AV block	85	13.1
Sinus node dysfunction	397	61.4
Other	41	6.3
Arrhythmia-related medication		
Class I	89	13.8
Class III	251	38.8
β-blocker	212	32.8
Calcium channel blocker	56	8.7
Digoxin	144	22.3

AF indicates atrial fibrillation; AV, atrioventricular; CABG, coronary artery bypass grafting; MVR, mitral valve replacement/repair; AVR, aortic valve replacement/repair; TVR, tricuspid valve replacement/repair; PVR, pulmonary valve replacement; and NYHA, New York Heart Association.

monitored days intersected with the monitored days of previously selected elements. AF was deemed to have been successfully identified if it was observed in at least 1 of the *n* sets of monitored days. This was performed for all patients of the study population, for monitoring durations of *k*=1, 2, 7, 14, and 30 days, and for strategies

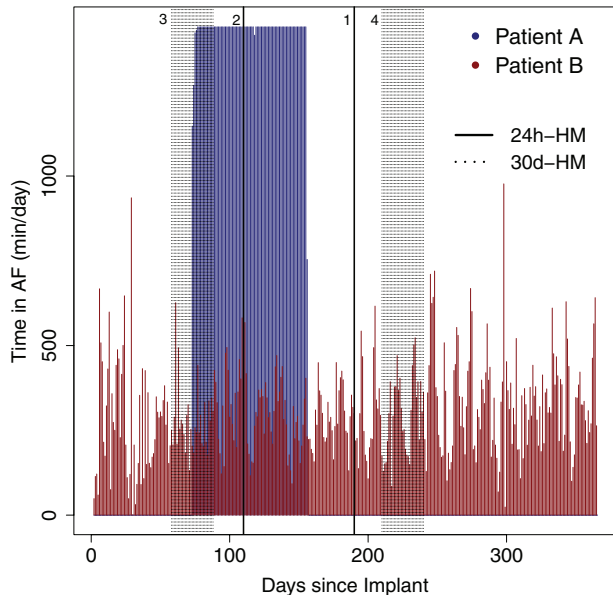


Figure 1. Reconstruction of the rhythm history of patients A and B. Both patients have similar atrial fibrillation (AF) burden (AF burden for patient A=0.22; AF burden for patient B=0.21). The vastly different temporal aggregation of the AF episodes and AF burden in these patients is responsible for the different sensitivities of intermittent AF monitoring modalities in these patients. For patient A, the probability of identification of AF recurrence in a 24-hour and 30-day monitor is $\text{Pr}(24\text{h}|\text{A})=0.23$ and $\text{Pr}(30\text{d}|\text{A})=0.26$. In this case, a random 30-day monitoring increases the sensitivity by only 3% from that which random 24-hour Holter monitoring (HM) would provide. For patient B, the respective probabilities are $\text{Pr}(24\text{h}|\text{B})=1$ and $\text{Pr}(30\text{d}|\text{B})=1$. The numbers 1 and 2 denote the time points of 2 random 24-hour HM tests, whereas the numbers 3 and 4 denote the time points of 2 random 30-day HM tests.

of $n=1, 2, 3, \dots, 12$ monitorings. The simulations were performed enough times ($>50\,000$) to allow stabilization of the inferred parameter. All simulated IRM was of the continuous recording type. Randomization tests were used to compare sensitivities of prolonged IRM strategies with shorter IRM strategies and to compare patients with high and low AF densities. The P values of 2-sided tests are reported.

Using computationally intensive simulation, we sought to evaluate the inferences that can be drawn from a series of negative IRM on the patients' AF burden. Upper 90% confidence intervals for the AF burden of patients with a series of negative examinations for a given IRM strategy were found by performing Monte Carlo simulations of the strategy and evaluating the 90th percentile of AF burden among those patients with a series of negative examinations in their simulated monitorings.

With the complete rhythm history of every patient reconstructed, we evaluated measures to describe the AF recurrence and burden development over time, taking into consideration the temporal aggregation of the episodes and the temporal dispersion of the AF burden over the monitored period. For each patient, the course of the AF burden development over time throughout the monitored period was analyzed, and the *minimum* monitored time required for the development of each proportion of the patient's total observed AF burden throughout the monitored period was calculated (blue or red dotted line in Figure 2, bottom). For example, in Figure 2, patient C develops 50% of his total burden in 11% of the monitored time (black dot, patient C). Patient D, in contrast, required 40% of the observation time to develop 50% of his burden (black dot, patient D) as each day contributes less to the total burden because the AF burden is spread over more days. The black diagonal line (Figure 2, patients C and D) represents a hypothetical uniform AF burden

development (AF burden spread evenly throughout the monitoring period). As a measure of temporal AF burden aggregation, the cumulative deviation of the minimum monitored time required for the development of each proportion of the patient's total observed AF burden throughout the monitored period from a hypothetical uniform burden development (area between the actual [dotted line] and the uniform burden development curves [solid diagonal], Figure 2) was calculated (blue or red highlighted area, Figure 2, bottom). This measurement was then scaled relative to the maximum possible burden aggregation for that specific patient's burden (ie, the complete burden as 1 continuous episode ["block of AF"]) to derive the burden density, an index taking values between 0 (meaning AF burden spread evenly over the observation time) and 1 (maximum possible AF burden aggregation [ie, "1 block/episode of AF"]).

For the numeric evaluation of the patient's AF density, we use the following definitions: For a patient with a total AF burden b , denote the minimum proportion of contiguous monitored time required for the development of a proportion p of the patient's total observed burden b as $F(p;b)$. The absolute cumulative deviation of the patient's actual burden development from the hypothetical uniform burden development is evaluated as $\int_0^1 |F(p;b) - p| dp$. When the burden b occurs with maximum temporal aggregation (ie, the complete burden as 1 continuous AF episode), the minimum time required for the development of 100% of the total burden is numerically equal to b , and the cumulative deviation from the hypothetical uniform burden development is $\frac{1-b}{2}$. AF density is numerically evaluated as follows:

$$\text{AF density} = 2 * \frac{\int_0^1 |F(p;b) - p| dp}{1-b}$$

The detection probability gain of a CM versus an IRM strategy of k duration [with a probability of AF detection in patient i of $\text{Pr}(k|i)$] was evaluated as $\frac{1 - \text{Pr}(k|i)}{\text{Pr}(k|i)}$ and depicts the increase in probability of AF detection that CM offers in comparison to that of a random IRM of k duration (k : 24 hours, 7 days, 14 days, 30 days). Linear regression was used to determine whether AF burden and AF density were independently associated with the probability gain. To restore normality, the log of the probability gain was used as dependent variable.

The Mann-Whitney test was used for comparing AF burdens among patients with high and low density. All statistical analyses and procedures were performed with R version 2.14.1 (R Development Core Team; 2011; <http://www.R-project.org/>).

Results

Sensitivity of IRM

Quantitative AF characteristics are displayed in Figure 3. Figure 4 displays the results for the sensitivity of random 24-hour and 7-, 14-, and 30-day monitorings at various frequencies to detect AF recurrence for the entire patient population. Prolonged IRM was significantly superior to shorter IRM ($P < 0.0001$ for all comparisons). However, a four 24-hour HM strategy had a sensitivity of only 52%, thus failing to identify AF recurrence in almost half of the patients with proven AF recurrence. Even with a theoretical strategy of three 30-day HM tests (for a total monitoring duration of 90 days per year), the sensitivity did not exceed 82%, and thus a nonnegligible proportion of patients with AF would be misclassified (Figure 4). However, the increased sensitivity offered by prolonged IRM durations comes at a cost of disproportionate increase in the required monitored time to achieve that level of sensitivity. For a sensitivity of 0.65, 30 days of monitoring would have been required with a 30-day

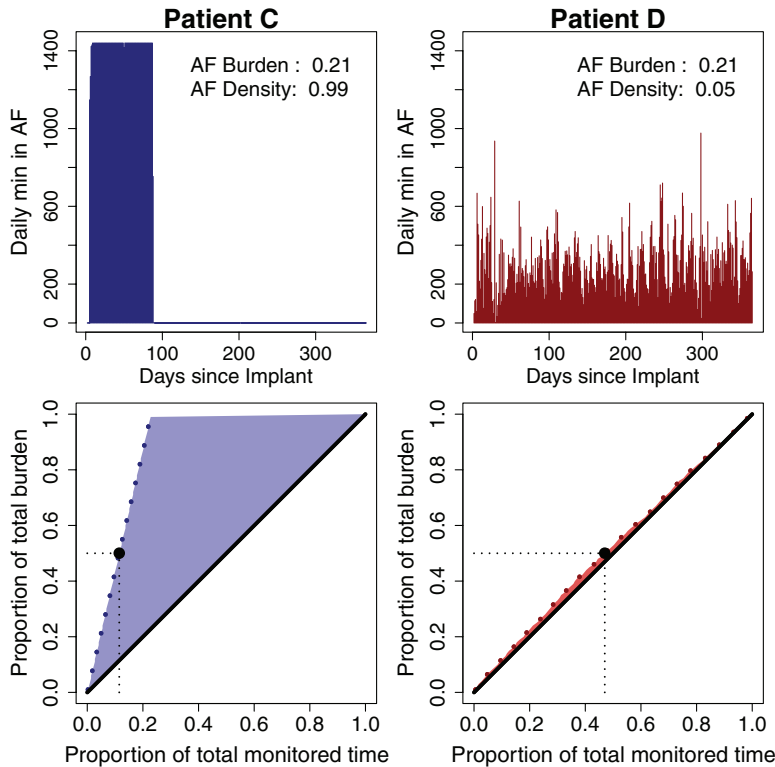


Figure 2. Two patients with the same atrial fibrillation (AF) burden but different burden aggregation (patient C, high-density AF; patient D, low-density AF). With the rhythm history reconstructed, the course of the minimum monitored time required for each burden proportion is plotted against the proportion of the total burden (dotted line, patients C and D, bottom). Patient C developed 50% of his total burden in 11% of the monitored time (black dot, patient C, bottom). Patient D, in contrast, required 40% of the observation time to develop 50% of his burden (black dot, patient D, bottom) as each day contributes less to the total burden because the AF burden is spread over more days. The black diagonal line (patients C and D, bottom) represents a hypothetical uniform AF burden development. The area between the actual (blue or red dotted line) and the uniform hypothetical (solid black diagonal) AF burden development is evaluated as a measure of the temporal aggregation of the AF burden (AF burden density).

HM strategy (one 30-day HM test), whereas for the same sensitivity, 7 monitored days would have been required with a 24-hour HM strategy (seven 24-hour HM tests) (Figure 4, solid horizontal black line).

Inferences Drawn From a Series of Negative IRM

Using computationally intensive simulation, we sought to evaluate the inferences that can be drawn from a series of negative IRM on the patients’ AF burden. Although IRM strategies cannot determine the precise AF burden, reliable

inferences can be obtained (for example, a patient with 4 negative random 30-day HM tests is highly unlikely to have a burden of >80%). These probabilities were derived from our patient population, taking into consideration not only the amount of AF in each patient but also the temporal characteristics of AF, and are displayed in Figure 5. A series of 4 negative random 30-day HM tests suggests with 90% confidence that this patient’s AF burden is <17%. Similarly, 8 negative 24-hour HM tests would be required to achieve the same level of confidence regarding the patient’s AF burden. IRM of shorter duration requires a higher number of examinations to achieve the same level of confidence as IRM of longer durations (Figure 5).

Influence of AF Burden and AF Density on Effectiveness on IRM

Using the reconstructed rhythm histories of our patients, we attempted to evaluate the quantitative as well as temporal AF characteristics that drive the sensitivity differences between IRM of various durations and frequencies. Burden density was directly related to the sensitivity of the different monitoring durations (Figure 6). Even with similar AF burdens (burden of low-density group: mean, 0.12±0.11; median, 0.07; quartile 1, 0.01; quartile 3, 0.18; burden of high-density group: mean, 0.13±0.14, median, 0.06; quartile 1, 0.01; quartile 3, 0.19; P=0.36), patients with low-density AF (AF density <0.5) achieve higher sensitivities for AF recurrence detection compared with patients with high-density AF (AF density >0.5; P<0.0001 for all comparisons of all monitoring duration sensitivities [24-hour, 7-day, 14-day, 30-day HM] between high- and low-density groups). In patients with high-density AF, AF recurrence is much more difficult to

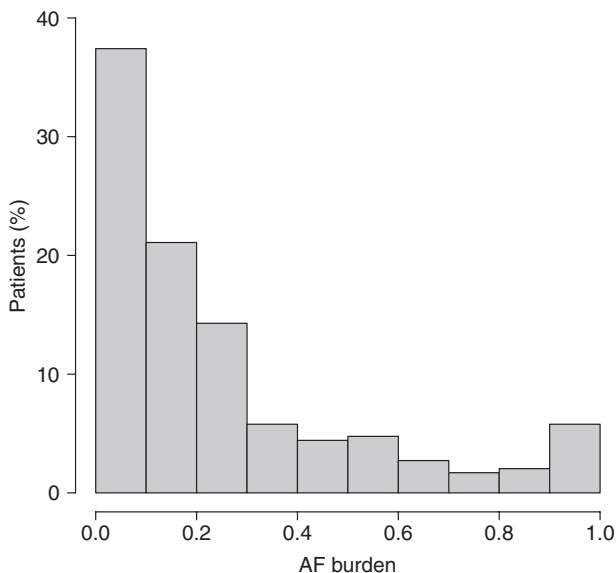


Figure 3. Histogram of the patients’ atrial fibrillation (AF) burdens. A total of 174 patients with no documented AF episode during the monitored time were excluded.

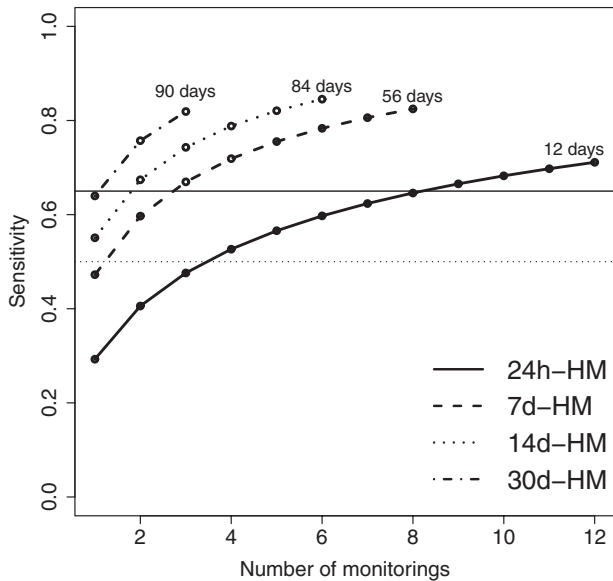


Figure 4. The sensitivity of intermittent rhythm monitoring (IRM) to detect atrial fibrillation (AF) recurrence as a function of monitoring frequency (number of random monitorings) and monitoring duration (24 hours and 7, 14, and 30 days) in patients with proven AF recurrence. Only patients with documented AF recurrence were included. Prolonged IRM was significantly superior to shorter IRM ($P < 0.0001$ for all comparisons). However, a four 24-hour Holter monitoring (HM) strategy had a sensitivity of only 52%, thus failing to identify AF recurrence in almost half of the patients with proven AF recurrence. The dotted horizontal line represents the sensitivity of 0.5 (fair coin). To achieve a sensitivity of 0.65 (solid horizontal line), 30 days of monitoring would have been required with a 30-day HM strategy (one 30-day HM test). To achieve the same sensitivity, 7 monitored days would have been required with a 24-hour HM strategy (seven 24-hour HM tests). The total number of monitored days required at the maximum frequency of each monitoring strategy is denoted.

capture with the use of IRM than in patients with low-density AF, even at similar AF burdens.

The superiority of CM in comparison to IRM was found to be significantly related to the AF burden and the AF density (Table 2). When CM is compared with 24-hour HM, the lower the AF burden and the higher the AF burden density, the greater is the gain from CM for AF recurrence detection (Figure 7, left). Patients with low-burden, high-density AF profit the most from CM. When CM was compared with a 30-day HM strategy, the gain obtained with CM was also found to be significantly related to the AF density and the AF burden. The higher the AF density and the lower the AF burden, the higher is the detection probability gain from the CM strategy (Figure 7, right).

Discussion

This study provides insights into the sensitivity of various monitoring strategies for the detection of AF recurrence with the use of data from a large patient population monitored with CM devices. Although 2 different monitoring devices were employed, AF detection in both devices was similar. The aim of our study was not to evaluate the AF detection process of the devices, which has been extensively validated previously,^{10,17–19} but to evaluate the characteristics and dynamics of AF recurrence detection after AF episode registration.

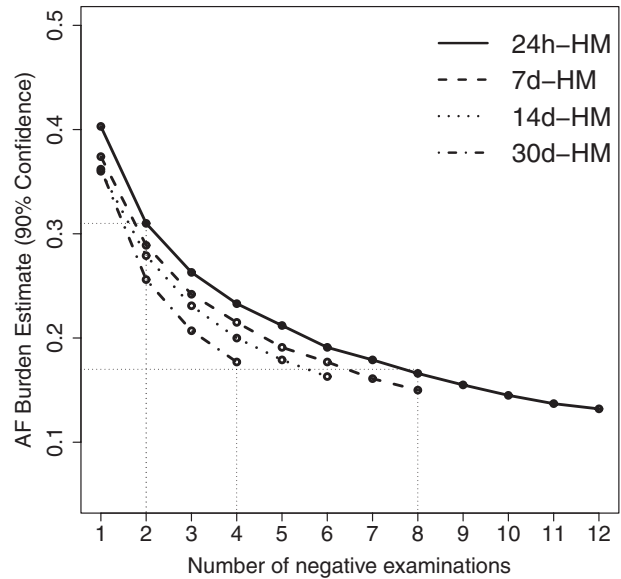


Figure 5. Inferences on atrial fibrillation (AF) burden drawn by hypothetical negative intermittent rhythm monitoring (IRM) examinations of the overall study population. For the study population, a series of 2 negative random 24-hour Holter monitor (HM) tests imply with a confidence of 90% that the patient's burden is $\leq 32\%$. Similarly, a series of 4 negative random 30-day HM tests imply with a confidence of 90% a burden of $\leq 17\%$. Eight negative series of 24-hour HM tests would have been required to obtain the same level of AF burden confidence.

Although this patient population presents an inhomogeneous collection of patients being monitored continuously with implantable devices, this gives us the opportunity to observe the characteristics of a very wide range of AF burden spectrum. The lower end of the AF burden spectrum is predominantly occupied by patients implanted with a CM device for monitoring after ablation procedures (mean AF burden, 0.05 ± 0.16 ; median, 0.002; quartile 1, 0.001; quartile 3, 0.01) or patients with paroxysmal, asymptomatic AF, whereas the higher end of the spectrum is predominantly occupied by patients implanted with a pacemaker device for symptomatic bradyarrhythmia and long-lasting persistent AF. This wide range of AF burden observed allowed us to evaluate the sensitivity of the IRM strategies and evaluate the AF burden dynamics over a wide range of AF burdens.

Judgment of Therapeutic Success Based on IRM

Our group and other groups have previously evaluated sensitivity differences between IRM and CM by comparing the information obtained from CM devices with actual IRM performed in the same patients.^{1,4,8,11} However, these methods have limitations because the time point of the IRM plays a major role in the evaluation of these methods. As depicted in Figure 1, in patient A, a 24-hour HM test performed at time point 1 will be negative and patient A will be regarded as “AF free,” whereas the same examination performed some days earlier at time point 2 would have identified AF recurrence. The results of such studies comparing continuous versus actual IRM are severely influenced by the choice of time points when the IRM is performed. Because in most of these studies only a limited number of IRM tests were compared

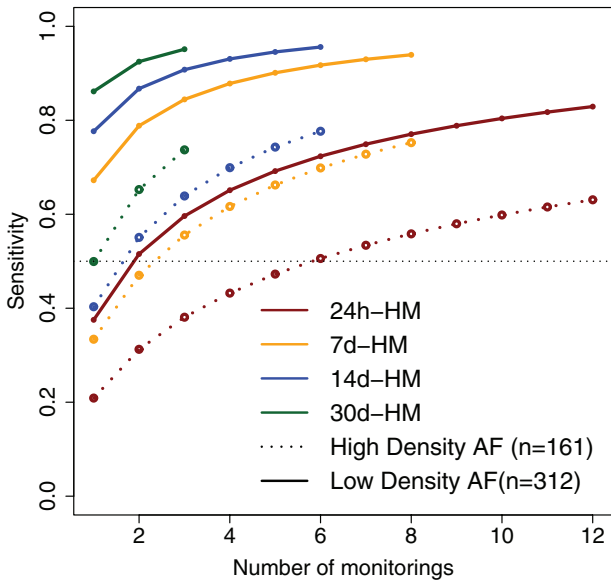


Figure 6. Sensitivity of various intermittent rhythm monitoring durations at various frequencies in patients with high (>0.5) and low (<0.5) atrial fibrillation (AF) burden density. Both high- and low-density AF groups had similar AF burdens (burden of low-density group: mean, 0.12±0.11; median, 0.07; quartile 1, 0.01; quartile 3, 0.18; burden of high-density group: mean, 0.13±0.14; median, 0.06; quartile 1, 0.01; quartile 3, 0.19; P=0.36). Low-density AF patients achieve higher AF recurrence detection sensitivity with less frequent monitoring than patients with high-density AF. In patients with high-density AF (blocks of AF), a significantly (P<0.0001) higher monitoring frequency is required to detect recurrence of AF. Only patients with documented AF recurrence were included. HM indicates Holter monitoring.

against CM, chance has an immeasurable effect on the results of such studies, and any inferences drawn from these results will be problematic. In the present study, the reconstruction of the rhythm history of every patient allows us to evaluate the true probability of AF recurrence detection of any random IRM of any duration in every patient, and computationally intensive simulation can draw inferences on the sensitivities and success rates of IRM of any duration and frequency.

Figure 4 presents the results of this approach in our patient collective, depicting the sensitivities of the 4 most common IRM durations (24-hour and 7-, 14-, and 30-day HM) at various monitoring frequencies. In our patient population, a monitoring strategy with four 24-hour HM tests would have a sensitivity of only 52%, thus failing to identify recurrence of AF in almost half of these patients. This low sensitivity becomes even more striking if one considers that the flipping of a fair coin would have provided the same sensitivity for the detection of AF recurrence in these patients. The direct implication of these results is that the evaluation of pharmacological or interventional therapies when a 24-hour HM is employed as a monitoring strategy should be interpreted cautiously because a great proportion of patients will be misclassified, AF recurrence will be underdiagnosed, and therapeutic success will be overestimated. Similarly, a single, random 30-day HM in our patient population will have a sensitivity of 63%, and thus the evaluation of the success of an ablation procedure with the use of a single, postinterven-

Table 2. Linear Regression Analysis of Effect of Burden and Burden Density on the Probability Gain of AF Recurrence Detection of 30-Day and 24-Hour HM vs Continuous Monitoring

Factor	Coefficient	P
Gain in AF recurrence detection: continuous vs 24-h HM		
AF burden	-5.76	<0.0001
AF burden density	2.61	<0.0001
Gain in AF recurrence detection: continuous vs 30-d HM		
AF burden	-4.08	<0.0001
AF burden density	4.84	<0.0001

AF indicates atrial fibrillation; HM, Holter monitoring. To restore normality, the log of the probability gain was used as dependent variable.

tional, 30-day HM would have a sensitivity of only 63% (slightly better than the sensitivity the flipping of a fair coin would provide). Even in the theoretical scenario of using three 30-day HM tests (for a total of 90 monitored days per year) would have a sensitivity of 82%; however, such a strategy seems unrealistic because it has been shown that complex or prolonged IRM strategies severely affect patient compliance.⁷ To obtain the sensitivities of the various IRM strategies presented in this report, a 100% patient compliance is assumed; however, this is rarely the case in everyday clinical practice because even short-duration IRMs pose a significant compromise in the patients' quality of life, and this has been shown to lead to monitoring discontinuation.^{7,20} Roten et al²⁰ recently showed that even with medium-duration IRM (7-day HM), 42% of patients complained of discomfort and skin irritation, whereas 16% discontinued the monitoring because of this reason. On the other hand, the implantation of leadless CM requires a minor surgical procedure, and the device itself may carry a risk of infection or patient discomfort. Prolonged IRM strategies not only have reduced patient compliance but also offer diminished returns in terms of sensitivity gained per monitored day. The latter problem is depicted in Figure 4. To achieve the same sensitivity (0.63; solid black horizontal line, Figure 4), 30 days of monitoring would have been required with a 30-day HM strategy (one 30-day HM test) versus only 7 days of monitoring with a 24-hour HM strategy (seven 24-hour HM tests). This disproportionate increase in required monitoring time to achieve a certain level of sensitivity between long- and short-duration IRM would certainly have an impact on patient compliance with aggressive IRM strategies.

If the AF burden temporal characteristics are also taken into consideration (Figure 6), the detection of AF recurrence in patients with high-density AF becomes even more problematic. In the case of patients with high AF burden densities, a strategy of six 24-hour HM tests or one 30-day HM test would have been needed to provide a sensitivity of 50% (similar to that of a fair coin). In the low-density group, six 24-hour HM tests or one 30-day HM test would have provided a sensitivity of 0.72 and 0.86, respectively. Thus, patients with high-density AF require an even more aggressive IRM strategy to detect AF recurrence and evaluate results of therapeutic interventions. It becomes clear that with

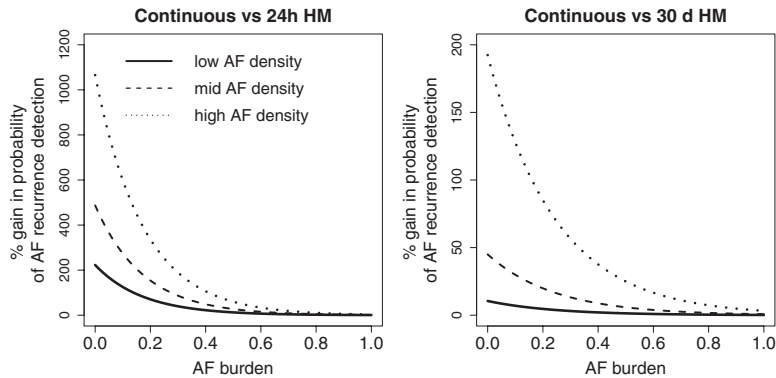


Figure 7. Gain in atrial fibrillation (AF) recurrence detection when using continuous monitoring versus 24-hour Holter monitoring (HM) (left) or 30-day HM (right) with respect to AF burden and burden density. Continuous monitoring offers significantly increased sensitivity in AF recurrence detection with a lower AF burden and a higher AF density (Table 2). The gain obtained with continuous monitoring diminishes at burdens >0.4 . AF burden density has been categorized into 3 categories (high, 0.6–1; mid, 0.3–0.6; low, 0–0.3).

intermittent monitoring altogether, a great part or even the majority of patients will be misclassified, and therefore the scientific evaluation of therapeutic interventions for AF becomes problematic.

Semiquantitative Estimation of AF Burden Based on Negative IRM

Although IRM may fail to detect AF recurrence in a great proportion of patients and cannot provide the level of information a CM strategy does, results of serial IRM may guide patient management in a semiquantitative way when CM is not possible because of either limited availability, patient compliance, or cost considerations. Such an evaluation is depicted in Figure 5. A series of negative IRM, although may underestimate qualitative AF detection per se, can however guide estimation of the potential AF burden level in patients prone to AF recurrence such as patients after ablation procedures, or patients on rhythm control strategies. In our large population of CM patients, 4 random negative 24-hour HM tests indicate, with a confidence of 90%, an AF burden of up to 23%. A higher number of serial negative IRM tests or negative IRM tests of greater duration can pinpoint the potential AF burden to lower levels. However, it is important to note that even average burdens that low have been shown to be a risk factor for thromboembolic complications.^{13–16,21}

AF Characteristics Influencing the Success of IRM

A major variable driving the sensitivity in detecting AF recurrence is not only the AF burden but the dispersion of it over time. AF density, as a measure of this dispersion, proved to be a major factor that influences the sensitivity of IRM at various frequencies and durations. Although patients A and B of Figure 1 have very similar amounts of AF burden over the same observation time, the probability of AF recurrence detection is vastly different because of the different temporal AF burden dispersion. In patient B (Figure 1), because of the even dispersion of AF burden, prolonged-duration IRM (30-day HM) is not superior to shorter-duration IRM (24-hour HM), the probability of AF recurrence detection of both IRM tests is 1, and the sensitivity of both strategies is 100%. In patient A (Figure 1), however, because of the high AF density, a random 30-day HM test offers only a negligible 3% increase in probability of AF recurrence detection compared

with a random 24-hour HM test. In the whole patient cohort (Figure 6), the AF density appears to have a major and statistically significant influence on the sensitivity of AF detection with the use of IRM strategies. In patients with high-density AF, the probabilities of AF recurrence detection and the sensitivities of IRM strategies are significantly inferior to those of patients with same level of AF burden but low AF density (Figure 6).

Implications for Clinical Outcomes, Patient Management, and Current Knowledge

Although the level of AF burden has been shown to be a significant risk factor for thromboembolic events,^{13–16} to the best of our knowledge, no study has taken into consideration the temporal characteristics of AF burden development. It is conceivable that the temporal distribution of AF burden may also play a role in the general thromboembolic risk that patients with AF face. Patients A and B (Figure 1), although spending the same time in AF, may have a different thromboembolic risk because of the different temporal distribution of their AF burden. Larger studies are required to evaluate this hypothesis, and prospective studies are required to evaluate the clinical impact of the combined quantitative and temporal AF characteristics on patient outcome.

The data presented may have some clinical implications. The results of our study depict the major limitations of IRM not only for management of patients but also for evaluation of the results of pharmacological as well as interventional therapies for AF. From a clinical perspective, accurate patient management can only be obtained with confidence with CM. From a scientific perspective, our evaluation until now of the success of pharmacological and interventional procedures based on IRM seems problematic in light of the results presented in this study. The accurate evaluation of therapies for AF in the era of evidence-based medicine mandates the use of CM.

It is conceivable that the widespread use of CM or the results from large studies of CM patients may require us to reevaluate our definitions of success or failure of therapeutic intervention for AF. It may very well be that the current definition of success of a therapeutic intervention only in terms of complete absence of AF recurrence might be inadequate and that a quantitative approach to AF recurrence may in the future require us to critically revisit our current knowledge and reevaluate our therapeutic interventions and

strategies, our results, and our patient management decision-making process. Whether the use of CM to guide clinical management can also improve patient outcomes remains uncertain at this point and must be investigated in large randomized controlled prospective studies.

Limitations

This study was underpowered to investigate the clinical impact of temporal AF characteristics on patient outcomes. We believe that this should be the scope of large prospective studies or registries of CM patients over sufficient follow-up time. AF episodes of patients primarily implanted with a CM device because of atrial tachycardia or atrial flutter have been included in the analysis. In our experience, although patients are typically classified in 1 distinct rhythm type (atrial flutter, AF, or atrial tachycardia), CM reveals that these patients actually tend to experience multiple types of atrial rhythms. Although 2 different devices were employed in the population of this study, previous studies have extensively investigated the accuracy of AF detection of these devices (sensitivity, $\geq 98\%$; specificity, $\geq 91\%$).^{10,17,18} Additionally, we tried to manually inspect all available episodes and, in cases of uncertainty or dispute, we discarded isolated, short (<5 minutes) episodes that could represent artifacts.

Conclusion

Patient follow-up with either short- or long-duration IRM is significantly inferior to CM for the detection of AF recurrence. IRM monitoring strategies will not identify AF recurrence in a great proportion of patients at risk. With short-duration, low-frequency IRM strategies, chance has an immeasurable effect on the evaluation of AF recurrence. Thus, results of pharmacological or invasive interventions evaluated with IRM should be considered with caution. Temporal AF characteristics play a major role in AF recurrence detection with the use of IRM strategies and may influence clinical parameters and outcomes in AF patients. For the scientific, evidence-based evaluation of AF treatments, CM should be strongly recommended.

Acknowledgments

We are thankful to Katrin Meyer and Martina Schröder for secretarial support and to Jana Paise, Anja Paap, and Tobias Frin for data documentation support.

Disclosures

Drs Charitos, Stierle, Baldewig, Robinson, and Sievers have no conflict of interest to disclose. Paul D. Ziegler is a full-time employee of Medtronic (>10 000 USD) and stockholder of Medtronic (>10 000 USD). Dr Hanke discloses modest lecture honoraria (<10 000 USD) from Medtronic.

References

- Hanke T, Charitos EI, Stierle U, Karluss A, Kraatz E, Graf B, Hagemann A, Misfeld M, Sievers HH. Twenty-four-hour Holter monitor follow-up does not provide accurate heart rhythm status after surgical atrial fibrillation ablation therapy: up to 12 months' experience with a novel permanently implantable heart rhythm monitor device. *Circulation*. 2009;120:S177–S184.
- Strickberger SA, Ip J, Saksena S, Curry K, Bahnson TD, Ziegler PD. Relationship between atrial tachyarrhythmias and symptoms. *Heart Rhythm*. 2005;2:125–131.
- Senatore G, Stabile G, Bertaglia E, Donnici G, De Simone A, Zoppo F, Turco P, Pascotto P, Fazzari M. Role of transtelephonic electrocardiographic monitoring in detecting short-term arrhythmia recurrences after radiofrequency ablation in patients with atrial fibrillation. *J Am Coll Cardiol*. 2005;45:873–876.
- Kottkamp H, Tanner H, Kobza R, Schirdewahn P, Dorszewski A, Gerdts-Li J-H, Carbucicchio C, Piorkowski C, Hindricks G. Time courses and quantitative analysis of atrial fibrillation episode number and duration after circular plus linear left atrial lesions: trigger elimination or substrate modification: early or delayed cure? *J Am Coll Cardiol*. 2004;44:869–877.
- Dubner S, Auricchio A, Steinberg JS, Vardas P, Stone P, Brugada J, Piotrowicz R, Hayes DL, Kirchhof P, Breithardt G, Zareba W, Schuger C, Aktas MK, Chudzik M, Mittal S, Varma N. ISHNE/EHRA expert consensus on remote monitoring of cardiovascular implantable electronic devices (CIEDs). *Europace*. 2012;14:278–293.
- Wilkoff BL, Auricchio A, Brugada J, Cowie M, Ellenbogen KA, Gillis AM, Hayes DL, Howlett JG, Kautzner J, Love CJ, Morgan JM, Priori SG, Reynolds DW, Schoenfeld MH, Vardas PE. HRS/EHRA expert consensus on the monitoring of cardiovascular implantable electronic devices (CIEDs): description of techniques, indications, personnel, frequency and ethical considerations. *Heart Rhythm*. 2008;5:907–925.
- Calkins H, Brugada J, Packer DL, Cappato R, Chen S-A, Crijns HJG, Damiano RJ Jr, Davies DW, Haines DE, Haissaguerre M, Iesaka Y, Jackman W, Jais P, Kottkamp H, Kuck KH, Lindsay BD, Marchlinski FE, McCarthy PM, Mont JL, Morady F, Nademanee K, Natale A, Pappone C, Prystowsky E, Raviele A, Ruskin JN, Shemin RJ. HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: recommendations for personnel, policy, procedures and follow-up: a report of the Heart Rhythm Society (HRS) Task Force on Catheter and Surgical Ablation of Atrial Fibrillation developed in partnership with the European Heart Rhythm Association (EHRA) and the European Cardiac Arrhythmia Society (ECAS); in collaboration with the American College of Cardiology (ACC), American Heart Association (AHA), and the Society of Thoracic Surgeons (STS): endorsed and approved by the governing bodies of the American College of Cardiology, the American Heart Association, the European Cardiac Arrhythmia Society, the European Heart Rhythm Association, the Society of Thoracic Surgeons, and the Heart Rhythm Society. *Europace*. 2007;9:335–379.
- Ziegler PD, Koehler JL, Mehra R. Comparison of continuous versus intermittent monitoring of atrial arrhythmias. *Heart Rhythm*. 2006;3:1445–1452.
- Pokushalov E, Romanov A, Corbucci G, Artyomenko S, Turov A, Shirokova N, Karaskov A. Use of an implantable monitor to detect arrhythmia recurrences and select patients for early repeat catheter ablation for atrial fibrillation: a pilot study. *Circ Arrhythm Electrophysiol*. 2011;4:823–831.
- Hindricks G, Pokushalov E, Urban L, Taborsky M, Kuck K-H, Lebedev D, Rieger G, Pürerfellner H. Performance of a new leadless implantable cardiac monitor in detecting and quantifying atrial fibrillation: results of the XPECT trial. *Circ Arrhythm Electrophysiol*. 2010;3:141–147.
- Edgerton JR, Mahoney C, Mack MJ, Roper K, Herbert MA. Long-term monitoring after surgical ablation for atrial fibrillation: how much is enough? *J Thorac Cardiovasc Surg*. 2011;142:162–165.
- Hindricks G, Piorkowski C, Tanner H, Kobza R, Gerdts-Li J-H, Carbucicchio C, Kottkamp H. Perception of atrial fibrillation before and after radiofrequency catheter ablation: relevance of asymptomatic arrhythmia recurrence. *Circulation*. 2005;112:307–313.
- Ziegler PD, Glotzer TV, Daoud EG, Wyse DG, Singer DE, Ezekowitz MD, Koehler JL, Hilker CE. Incidence of newly detected atrial arrhythmias via implantable devices in patients with a history of thromboembolic events. *Stroke*. 2010;41:256–260.
- Glotzer TV, Daoud EG, Wyse DG, Singer DE, Ezekowitz MD, Hilker C, Miller C, Qi D, Ziegler PD. The relationship between daily atrial tachyarrhythmia burden from implantable device diagnostics and stroke risk: the TRENDS study. *Circ Arrhythm Electrophysiol*. 2009;2:474–480.
- Daoud EG, Glotzer TV, Wyse DG, Ezekowitz MD, Hilker C, Koehler J, Ziegler PD. Temporal relationship of atrial tachyarrhythmias, cerebrovascular events, and systemic emboli based on stored device data: a subgroup analysis of TRENDS. *Heart Rhythm*. 2011;8:1416–1423.

16. Healey JS, Connolly SJ, Gold MR, Israel CW, Van Gelder IC, Capucci A, Lau CP, Fain E, Yang S, Bailleul C, Morillo CA, Carlson M, Themeles E, Kaufman ES, Hohnloser SH. Subclinical atrial fibrillation and the risk of stroke. *N Engl J Med*. 2012;366:120–129.
17. Sarkar S, Ritscher D, Mehra R. A detector for a chronic implantable atrial tachyarrhythmia monitor. *IEEE Trans Biomed Eng*. 2008;55:1219–1224.
18. Swerdlow CD, Schols W, Dijkman B, Jung W, Sheth NV, Olson WH, Gunderson BD; for the Worldwide Jewel AF Investigators. Detection of atrial fibrillation and flutter by a dual-chamber implantable cardioverter-defibrillator. *Circulation*. 2000;101:878–885.
19. Purerfellner H, Gillis AM, Holbrook R, Hettrick DA. Accuracy of atrial tachyarrhythmia detection in implantable devices with arrhythmia therapies. *Pacing Clin Electrophysiol*. 2004;27:983–992.
20. Roten L, Schilling M, Häberlin A, Seiler J, Schwick NG, Fuhrer J, Delacrétaz E, Tanner H. Is 7-day event triggered ECG recording equivalent to 7-day Holter ECG recording for atrial fibrillation screening? *Heart*. 2012;98:645–649.
21. Hart RG, Pearce LA, Rothbart RM, McAnulty JH, Asinger RW, Halperin JL; Stroke Prevention in Atrial Fibrillation Investigators. Stroke with intermittent atrial fibrillation: incidence and predictors during aspirin therapy. *J Am Coll Cardiol*. 2000;35:183–187.

CLINICAL PERSPECTIVE

Patient follow-up with either short- or long-duration intermittent rhythm monitoring is significantly inferior to that with continuous rhythm monitoring for the detection of atrial fibrillation (AF) recurrence. Intermittent rhythm monitoring strategies fail to identify AF recurrence in a great proportion of patients at risk. Even with feasible, aggressive intermittent monitoring strategies (such as quarterly 24-hour Holter monitoring), AF recurrence will not be identified in a significant proportion of patients. Because of the low sensitivity of intermittent rhythm monitoring, the results of pharmacological or invasive interventions for AF when evaluated with intermittent monitoring should be considered with caution because chance has an immeasurable effect on the detection of AF recurrence, and thus a great proportion of patients will be misclassified, AF recurrence will be underdiagnosed, and therapeutic success will be overestimated. Novel quantitative and temporal AF characteristics (AF burden, AF density) play a major role in AF recurrence detection with the use of intermittent monitoring strategies and may influence clinical parameters and outcomes in AF patients. For accurate patient management, as well as for the scientific, evidence-based evaluation of AF treatments, continuous rhythm monitoring should be strongly recommended.

Go to <http://cme.ahajournals.org> to take the CME quiz for this article.