MEDPAGE TODAY®

Blogs > Revolution and Revelation

What Did the Apple Heart Study Really Find?

— Milton Packer dissects the data and proposes a formal challenge to tech giant

by Milton Packer MD

March 20, 2019

Six months ago in this space, I wondered if the new Apple Watch was the worst heart device ever.

Some readers may have thought I was exaggerating. Others may have imagined that I was being too cynical.

But based on the results of the Apple Heart Study -- which were released at the American College of Cardiology meeting this week -- it is now official. The Apple Watch is a serious competitor for the worst heart device ever.

What was the Apple Heart Study? It was one of the largest studies of "digital health" to date. Nearly 420,000 people throughout the U.S. agreed to participate. The participants wore watches that sent intermittent information about the regularity of their heart

rhythm to Apple for varying periods of time. During the study, about 2,100 notifications of a heart "irregularity" were sent.

Did the Apple Watch detect any serious cardiac arrhythmias in a meaningful number of people?

A majority (219,179, 52%) of the people in the study were under age 40. Among those people, a mere 341 (0.16%) were notified of an "irregularity," and of these, only nine (0.004%) actually had atrial fibrillation. Were 97% of those younger people receiving alerts scared unnecessarily?

What were the results in elderly people?

Of 24,626 people ages 65 or older, 3.14% were notified of an "irregularity," and of these, only 63 people (0.26%) actually had atrial fibrillation. But the device sent an alert to 775 people! Were 90% of elderly people who received an alert scared unnecessarily?

Overall, the chances of the Apple Watch detecting undiagnosed atrial fibrillation in this study were lower than the chance of a person being struck by lightning during their lifetime (0.03%)!

How long did these episodes of atrial fibrillation last? In 75%, the rhythm lasted for 6 hours or less. Physicians have no idea whether such brief episodes are clinically important.

How many episodes of atrial fibrillation were not detected by the device in the Apple Watch study? We do not know. The study did not collect those data.

But an analysis by Venk Murthy, MD, PhD, of the University of Michigan, suggests that the Apple device failed to correctly identify atrial fibrillation in a large proportion of people who actually had it.

Was the study a success?

The investigators planned a prespecified measure of precision to determine the success of the study. The threshold for success focused on people ages 65 or older and was set at a relatively easy target, especially given the very large size of the study. And yet, even that prespecified measure of success was NOT achieved in the Apple Heart Study. So the study was not a success according to its very own criterion.

What is the risk of death in asymptomatic individuals who have atrial fibrillation first

detected by an Apple Watch? We do not know.

What is the risk of stroke in asymptomatic individuals who have atrial fibrillation first detected by an Apple Watch? We do not know.

Can we prevent death or stroke in asymptomatic individuals who have atrial fibrillation first detected by an Apple Watch? We do not know.

Should people who have atrial fibrillation first detected by an Apple Watch receive long-term anticoagulation? We do not know.

How many people who receive anticoagulation for atrial fibrillation first detected by an Apple Watch experience serious bleeding resulting from the use of the anticoagulants? We do not know.

So more than 400,000 people participated in the Apple Watch study. Most of the 2,100 notifications of an irregularity were not confirmed as atrial fibrillation. An untold number of cardiac arrhythmias were not detected. And we learned almost nothing of importance.

How many people had their lives saved as a result of using the Apple Watch? As far as anyone can tell, the answer is zero.

How many people had strokes prevented as a result of using the Apple Watch? As far as anyone can tell, the answer is zero.

How many people were harmed as a result of the information they received? We do not know. But many people were made unnecessarily anxious for meaningful periods of time after they were informed of the possibility that they might have a heart problem that they did not have.

So let us go back 6 months -- to the date of my original blog post on this subject.

At that time, I wrote: "First, the device is likely to alert many truly healthy people, causing enormous and unnecessary anxiety. Most will be motivated to seek medical attention, potentially putting horrific pressures on our healthcare system. Even a low false positive error rate is problematic in a low-prevalence population, leading to meaningful squandering of healthcare resources."

I also wrote: "Second, the device might detect atrial fibrillation in a few people who did

not previously know that they had the arrhythmia. But what benefits does this knowledge provide? What does the presence of these asymptomatic bursts mean? Do they have clinical importance? Should they be treated? No one knows."

Shortly after the launch of the Apple Watch, Venk Murthy carried out an analysis that was based on the data that led the FDA to clear the device in the first place (based on <1,000 people). That analysis predicted the results of the 400,000+ person Apple Heart Study very closely.

So now, after a massive study, we know almost nothing new, as compared with the data available at the time of the Apple Watch launch. (Disclosure: In November 2018, I received a one-time consulting fee from Johnson & Johnson in connection with the Apple heart monitoring function.)

Are there any lessons learned from the Apple Heart Study?

Six months ago, Apple COO Jeff Williams dubbed the Watch the "ultimate guardian for your health." He was wrong.

Six months ago, at a heavily promoted media event, Dr. Ivor Benjamin, president of the American Heart Association, referred to the heart app as "game-changing."

Was he right? Well, it depends what game you are playing.

Dr. Lloyd Minor, Stanford's dean of medicine and a head-and-neck surgeon, thinks that the study "provides very encouraging evidence that a device, the Apple Watch, can be used to detect a-fib and to point out to people when additional monitoring or testing may be needed." My response to Dr. Minor: the evidence from the study is far from encouraging. The data actually reinforce all the concerns we had 6 months ago.

Apple paid Stanford University more than $8 million for the study. For Apple, this amount is too small to even count as a rounding error in its annual profits.

But for millions of people who are exposed to the company's powerful marketing hype, the price they pay may be far greater than the price of the watch.

How can Apple actually show that the heart monitoring device is really helpful to people who buy the watch?

Apple could sponsor a definitive trial to prove the worth of their heart monitor function

in people not known to have atrial fibrillation. In the trial, every participant would wear the device. Half would be randomized to receive information about heart irregularities from the watch, and half would not. People who are informed of an irregularity would be advised to seek medical attention. Every randomized person would be followed for the primary endpoint of the occurrence of death, disabling stroke, cardiac arrest, or serious bleeding (the CABANA endpoint) for at least one year. The trial would have sufficient power to detect a meaningful effect on the occurrence of stroke.

If the device really does any good, the trial would be able to show a real benefit of using the Apple Watch in a very reasonable period of time. Even if the trial focused on high-risk patients, it would need to be twice as large as the Apple Heart Study. But at least such a trial would provide critically useful information -- which is currently what no study to date has provided.

Is Apple going to do a study that is actually able to show whether its device is a "guardian of health" or a "game-changer"?

In writing this post, I am publicly asking Apple to do the right study. Make sure that everyone in the trial has a watch, but give only half of the participants information from the watch. Owners of the watch would be randomized into one of the two groups to ensure that they are truly comparable. And Apple must make sure that the study has sufficient power to detect a meaningful reduction in the risk of stroke without an offsetting increase in major bleeding. That is what the detection of atrial fibrillation is all about.

To all of my readers, please consider my proposal to be a formal challenge to the company.

Will Apple take up my challenge? We will see.

Packer recently consulted for Actavis, Akcea, Amgen, AstraZeneca, Boehringer Ingelheim, Cardiorentis, Daiichi Sankyo, Gilead, J&J, Novo Nordisk, Pfizer, Sanofi, Synthetic Biologics, and Takeda. He chairs the EMPEROR Executive Committee for trials of empagliflozin for the treatment of heart failure. He was previously the co-PI of the PARADIGM-HF trial and serves on the Steering Committee of the PARAGON

of the PARADIGM-HF trial and serves on the Steering Committee of the PARAGON-HF trial, but has no financial relationship with Novartis.