Robotic Therapy: The Tipping Point

H. I. Krebs\textsuperscript{1,2} and N. Hogan\textsuperscript{1,3}

\textsuperscript{1}Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge, MA, USA
\textsuperscript{2}University of Maryland School of Medicine, Department of Neurology and Division of Rehabilitative Medicine, Baltimore, MD, USA
\textsuperscript{3}Massachusetts Institute of Technology, Department of Brain and Cognitive Science, Cambridge, MA, USA

Abstract

The last two decades have seen a remarkable shift in the neuro-rehabilitation paradigm. Neuroscientists and clinicians moved away from the perception that the brain is static and hardwired, to a new dynamic understanding that plasticity is a fundamental property of the adult human brain and might be harnessed to remap or create new neural pathways. Capitalizing on this innovative understanding, we introduced a paradigm shift in the clinical practice in 1989 when we initiated the development of the MIT-Manus robot for neuro-rehabilitation and deployed it in the clinic in 1994. Since then, we and others have developed and tested a multitude of robotic devices for stroke, spinal cord injury, cerebral palsy, multiple sclerosis, and Parkinson’s disease. Here we discuss whether robotic therapy has achieved a level of maturity to justify its broad adoption in the clinical realm as a tool for motor recovery.

Keywords

Robotic Therapy; Assistive Technology

Disruptive Technology

Disruptive technology is a term coined to characterize an innovation that disrupts an existing market or way of doing things and creates a new value network. The concept was first described at Harvard Business School by Clayton M. Christensen, who described the concept in 1996 as: “Generally, disruptive innovations were technologically straightforward, consisting of off-the-shelf components put together in a product architecture that was often simpler than prior approaches. They offered less of what customers in established markets wanted and so could rarely be initially employed there. They offered a different package of
attributes valued only in emerging markets remote from, and unimportant to, the mainstream.” Eventually with improvement, borrowing from Malcolm Gladwell, the moment of critical mass, the threshold, the boiling point is reached and the old practices and existing value network is abandoned in favor of the new one. The tipping point occurs not because the existing network is not focused on the “customer;” on the contrary, it is, but sometimes too much focus on the customer can be counterproductive and inhibit new ideas.

Two good examples of disruptive technologies with significant impact in the Boston area have to do with personal computers and digital photography. Ken Olsen and Harlan Anderson were working at MIT and noticed that students would line up for hours waiting for interactive terminals. They built Digital Equipment Corporation (DEC) based upon this concept of interactive behavior between users and computer and created a powerhouse that in 30 years became the second largest computer company, threatening the hegemony of IBM (1957–2002). DEC continued to pay close attention to the customer as it developed the PDP family and the VAX mini-computer. After several market research studies, its founder, Ken Olsen, famously stated his opinion about personal computers in 1977 that “there is no reason for any individual to have a computer in his home.” By the time the tipping point had occurred in the early 1990s, it was too late for DEC and its days of glory (and existence) were over. The story of Polaroid has a DEC déjà-vu flavor. Polaroid was a synonym for instant film. Several market research studies led Polaroid management to ignore the effect of digital cameras on its film business. By mid-2000s, it was essentially liquidated (but for the brand name).

### Does Robotic Therapy Constitute a Disruptive Technology?

Capitalizing on the new understanding that the brain contains dynamic networks capable of remapping and creating new pathways following an injury, we introduced a paradigm shift in clinical practice in 1989 when we initiated the development of the MIT-Manus robot for neuro-rehabilitation under the sponsorship of the National Science Foundation (NSF) and, as defined by Christensen, offered a different package of attributes valued only in an emerging market of clinical visionaries remote from, and initially unimportant to, mainstream rehabilitation practices. Fletcher McDowell MD, the former CEO of the Burke Rehabilitation Hospital and Foundation, is one of these clinical visionaries in neuro-rehabilitation. Following NSF 5-year sponsorship to develop the technology, he sponsored its deployment in the clinic in 1994 and nurtured it for the next 5 years until NIH and VA funds were secured to continue the research and other groups around the country and world followed our footsteps. In what follows, we will argue that since 2010 the tipping point of robotic therapy was reached at least for the upper extremity (UE) robotic therapy, moving the field beyond clinical visionaries into the mainstream.

#### Upper Extremity Robotic Therapy: The Tipping Point

Since the publication of the first controlled study with stroke inpatients, several studies have been completed with both stroke inpatients and outpatients demonstrating the potential of robotic therapy for the upper extremity. These results were discussed in different meta-analysis (see for example) and led to the 2010 American Heart Association (AHA) guidelines for stroke care which recommended that: “Robot-assisted therapy offers the amount of motor practice needed to relearn motor skills with less therapist assistance… Most trials of robot-assisted motor rehabilitation concern the upper extremity (UE), with robotics for the lower extremity (LE) still in its infancy… Robot-assisted UE therapy.

---

1The number of famous CEOs who underestimated the impact of personal computers is quite remarkable. There is another famous quote attributed to Watson from IBM: "I think there is a world market for about five computers.”
however, can improve motor function during the inpatient period after stroke.” AHA suggested that robot-assisted therapy for the UE has already achieved Class I, Level of Evidence A for Stroke Care in the Outpatient Setting and Care in Chronic Care Settings. It suggested that robot-assisted therapy for UE has achieved Class IIa, Level of Evidence A for stroke care in the inpatient setting. Class I is defined as: “Benefit >>> Risk. Procedure/Treatment SHOULD be performed/administered;” Class IIa is defined as: “Benefit >> Risk, IT IS REASONABLE to perform procedure/administer treatment;” Level A is defined as “Multiple populations evaluated: Data derived from multiple randomized clinical trials or meta-analysis” 15.

This is not an isolated opinion. The 2010 Veterans Administration/Department of Defense guidelines for stroke care came to the same conclusion endorsing the use of rehabilitation robots for the upper extremity, but went further to recommend against the use of robotics for the lower extremity. More specifically, the VA/DOD 2010 guidelines for stroke care “Recommend robot-assisted movement therapy as an adjunct to conventional therapy in patients with deficits in arm function to improve motor skill at the joints trained.” For the lower extremity the VA/DOD states that “There is no sufficient evidence supporting use of robotic devices during gait training in patients post stroke.” The VA/DOD issued a recommendation to employ robot-assisted therapy for the UE: “A recommendation that clinicians provide (the service) to eligible patients. At least fair evidence was found that the intervention improves health outcomes and concludes that benefits outweigh harm.” Regarding the lower extremity, the VA/DOD suggested against robot-assisted therapy: “Recommendation is made against routinely providing the intervention to asymptomatic patients. At least fair evidence was found that the intervention is ineffective or that harms outweigh benefits” 14.

Subsequent to the publication of the AHA and VA/DOD guidelines, the results of the largest single study of robotics have become available. The multi-site, independently run, Veterans Affairs trial CSP-558 (VA-ROBOTICS) on chronic stroke of upper extremity robotics employing the commercial version of the MIT-Manus for shoulder-and-elbow therapy together with the corresponding anti-gravity, wrist, and hand robots 13.

VA-ROBOTICS included 127 Veterans with chronic stroke at least 6 months post-index stroke with an impairment level characterized by very severe to moderate (Fugl-Meyer Assessment between 7 to 38 out of 66 points for the upper extremity). Veterans with multiple strokes were included in this study that lasted for 36 weeks: a 12-week intervention followed by a follow-up period lasting 6-months. Veterans were randomly assigned to either the robotic therapy group (RT, N=49), the intensity-matched comparison group (ICT, N=50), and the usual care group (UC, N=28). VA-ROBOTICS compared the efficacy of robot-mediated therapy (RT) to usual care (UC) and to intensive comparison therapy (ICT). Usual Care was not dictated or prescribed by the protocol. The treatment was allowed to vary as per therapy targeting specifically the upper extremity, which consisted of an average of 3 sessions per week from therapists delivering treatment as they deemed clinically appropriate for the upper extremity. To encourage subject retention, subjects randomized to UC selected at trial completion their choice of an additional 12 weeks of robotic therapy or ICT. The RT group received 3 sessions per week of robotic training for the shoulder-and-elbow, wrist, and hand that delivered 1,024 movements per session. The ICT group received 3 sessions per week of a therapy created to have a therapist deliver comparable movement intensity and repetition as the RT group during the same period. Contrary to other studies of robot-assisted rehabilitation that employed a control intervention expected to have little effect on the primary outcome, VA-ROBOTICS is unique in that it included an active control treatment group. The study was based on the hypothesis that RT group would experience greater improvement in motor impairment at 12 weeks compared with the UC
and ICT groups, as measured by the upper extremity component of the Fugl-Meyer scale. Of note, the ICT intervention is not conventional therapy. It employs manual techniques but would likely be impractical to implement as clinical therapy. It is unlikely that therapists could consistently assist the paretic arm during standard clinical care for almost 1,000 movements per session as done for the ICT group (instead of the typical 45 movements per session in usual care for chronic stroke patients). We created this control treatment specifically to compare the treatment effect of the robot providing assistance with these movements versus a human therapist providing assistance with the same movements.

RESULTS

Safety

VA-ROBOTICS evaluated the safety of the employed rehabilitation robots. There were no serious adverse events in the RT group. A few patients mentioned muscle soreness, which is not surprising considering that they were making 1,024 movements in an hour robot session with the paretic limb.

Clinical Outcomes

The first and perhaps most understated finding of the VA-ROBOTICS was that usual care did not reduce impairment, disability, or improve quality of life in chronic stroke survivors. It is important to note that one-third of the usual care group received ongoing conventional upper-extremity therapy for an average of three sessions per week. The usual care intervention had no measurable impact and, to conserve financial resources, it was discontinued as futile midway through the study.

The comparison between the RT and UC groups included: a) the comparison between the robot group and usual care subjects which involved roughly only the first half of the RT group while the UC was not discontinued, and b) whether the changes were robust and long lasting. On this score, robot therapy was statistically superior to usual care in Stroke Impact Scale (quality of life) at the completion of the intervention and also in the Fugl-Meyer (impairment) and Wolf Motor Function (function) 6 months following the completion of the intervention.

The results are far more impressive if we compare the whole robot therapy group with the usual care and not just the analysis that focused on the first half of the study. While the results at 12 weeks showed that the difference between the first half of the robotic treatment group and usual care was over 2 Fugl-Meyer points, the difference between the second half of the robotic treatment group and usual care was almost 8 points in the Fugl-Meyer assessment with the comparison of the total robotic group versus the total usual care showing a 5-point change (see Figure 1).

The reason(s) for the smaller clinical effects of the robotic intervention in the first stage of the study when compared with the second stage of the study are debated. We contend that this discrepancy is most likely due to the omission of a “phase-in” stage in this study. When testing a new therapy, it is common practice to treat a predetermined number of subjects during the initial phase of the trial with the new therapy at each site before beginning data collection for the actual controlled trial in order to gain familiarity and expertise with the novel treatment. This hypothesized learning effect for therapists supervising robotic therapy is perhaps worthy of further study, as some might have expected robotic therapy to be operator independent. Nevertheless VA-ROBOTICS demonstrates the robustness of the results: even when therapists are learning how to use the novel tools, the results are better than usual care.
The comparison between the RT and ICT groups did not show any differences \textsuperscript{16}. That said, the study biostatisticians selected the most conservative approach to deal with the data. They employed a fixed-model instead of a mixed-model to estimate the 36 weeks’ outcome. They estimated a single model using data from all the patients in the study (UC, RT, ICT) and then employed this model to estimate the outcomes at follow-up for each intervention. Because the interventions are different, we argue that it may be appropriate to use a less conservative approach and employ mixed-models. Figure 2 illustrates the impact of fixed-model on the data. A direct measurement at 36 weeks demonstrates that the RT leads to an advantage of approximately 2 points in the Fugl-Meyer assessment. However, employing a fixed model depressed the measurement of the RT group at 36 weeks and pumped-up the ICT group. This leads to an estimate favoring the ICT over the RT group of $-0.58$ in Fugl-Meyer assessment. However, none of these differences were significant.

Note also that patients in the RT group continued to improve even after the intervention was completed at 12 weeks. Thus, the continued and persistent improvement at the 6-month follow-up evaluation suggests improved robustness and perhaps an incremental advantage that prompted further improvement even without intervention. For example, an improvement of roughly 3 points in the Fugl-Meyer scale might enable a very severe patient to start to raise his/her arm and to bathe independently, or to start to stretch the formerly paralyzed arm so that independent dressing could take place. It might enable a more moderate stroke patient to start to tuck in the shirt or to hike the pants independently, or to start to reach overhead and actively grasp an object.

This continued improvement after completion of the intervention is quite remarkable as VA-ROBOTICS included patients with chronic stroke disability in the moderate to severe range and over 30\% had multiple strokes. As such, the group represented a spectrum of disability burden that many studies have avoided and, in our case, represented the majority of the cases. Note that 65\% of the volunteers interviewed were enrolled. It suggests that robotic therapy for the upper extremity offers an opportunity to a broad spectrum of stroke patients.

**Cost Outcomes**

In this era of cost containment, an important and unexpected result arose from the recently completed cost-benefit analysis \textsuperscript{17}. The purchase cost of the four robotic modules (shoulder-elbow, wrist, anti-gravity, and hand) was $230,750; the interest rate on borrowing to purchase these robots was estimated at 6.015\% with 33\% facility overhead on top of the purchase value, and a $5,000 annual maintenance fee per robot. Yet, the additional cost of delivering RT or ICT was $5,152 and $7,382 respectively and the difference was statistically significant (P<0.001). While the active interventions (RT and ICT) added cost, when we compared total cost which includes the clinical care needed to take care of these Veterans for the 36 weeks of the trial (12 weeks of intervention and 6 months without any active intervention), there were no differences between active intervention and usual care. The total cost for the VA was roughly the same: $17,831 for robot therapy, $19,746 for the intensive comparison group, and $19,098 for the usual care. The usual care group used the rest of the health care system more often than the active intervention groups. In other words, for 36 weeks of care the robotic group cost the VA $5,152 for robotic therapy and $12,679 for clinical care. For 36 weeks of care the usual-care group cost the VA approximately $19,098.

We initially speculated that perhaps the surprising decreases in healthcare cost were due to a Hawthorne effect. The active groups were receiving extra attention during the 36 weeks’ trial duration. To determine whether that was the case, the VA health economists (Palo Alto VA, Stanford University, CA) continued to collect cost data on these patients. One may speculate that if placebo accounted for a significant component of these cost reductions, then costs should trickle up after trial completion. On the contrary, they did not for the robotic
therapy group. The health care cost until the end of September 2009 (after the 36 weeks of
the trial) averaged $7,777 for the RT group and $14,513 for the ICT group; this difference
was statistically significant (P<0.04).

One needs to take these results with the appropriate caveats--the sample size was small,
predominantly male, severe to moderate strokes, the data variability was large--nevertheless,
the results suggest better care for the same total cost.

**Lower Extremity Robotic Therapy: In Its Infancy**

The two most common lower extremity (LE) robotic rehabilitation devices are the Lokomat
(Hocoma, Switzerland) and the Autoambulator (Healthsouth / Motorika, Israel). The
estimates are that there are already around 350 Lokomats and around 100 Autoambulators in
clinical settings, yet the negative perception of LE robotic rehabilitation is not without merit.
While the installed robotic base is reasonably large, there are few published randomized
controlled studies supporting their use. In fact, some of the large studies employing the
Lokomat (Hocoma, Zurich, Switzerland) showed statistically significantly inferior results
when compared to those produced by usual care as practiced in the US for both chronic as
well as for sub-acute stroke patients. Of course, the characteristics and intensity of usual
care might vary according to the country and, hence, it is important to acknowledge that
these results comparing Lokomat training with usual care are primarily valid for the US.
However it is appropriate to note that Healthsouth, who first developed and has employed
the largest numbers of Autoambulators for almost 10 years, has not published any controlled
studies on that device and the VA, which employs the largest number of Lokomats, has
recently published a guideline recommending against its use post-stroke by its clinicians.

Figure 3 shows the outcomes of two studies comparing LE rehabilitation robotics with usual
care. The top row shows results with chronic stroke patients (stroke onset > 6 months), who
trained 3 times per week for 30 minutes for 4 weeks, demonstrating improvements for the
Lokomat trained (white bars) and the usual care group (black bars). However the usual care
group improved significantly more than the Lokomat trained group. Of note, the same
inferior results were observed for all levels of impairment, i.e., both severe and moderate
strokes appear to benefit more from usual care than Lokomat training. For sub-acute stroke
patients (stroke onset < 6 months) who trained for 8 weeks, a qualitatively similar result was
observed. Both groups improved from admission to mid-point, to completion, and to a 3
months’ follow-up but patients in the usual care group improved more and the difference
between groups was statistically significant.

There are many plausible reasons for these results and the apparent immaturity of lower
extremity robotic therapy. We would like to highlight that perhaps we need to better
understand the difference between “best practices” and tested practices. Clinicians and
technologists assumed that body-weight-supported treadmill training (BWSTT) delivered by
two or three therapists was an effective and superior form of therapy compared to usual care
and, thus, automating BWSTT appeared to be a logical approach. However, an NIH-
sponsored randomized controlled study (RCT) demonstrated that, contrary to the hypothesis
of its clinical proponents, body-weight-supported treadmill training administered by 2 or 3
therapists for 20 to 30 minutes followed by 20 to 30 minutes of overground carry-over
training did not lead to superior results when compared to a home program of strength
training and balance (LEAPS Study). These are landmark results that must be seriously
acknowledged by both roboticists and clinicians: the goal of rehabilitation robotics is to
optimize care and augment the potential of individual recovery. It is not to automate current
rehabilitation practices which for the most part lack a scientific evidential basis, primarily
due to the lack of tools to properly assess the practices themselves.
Defining Success in Robotic Rehabilitation

We must define the benchmark required to determine whether a disruptive technology has gone beyond a different package of attributes valued only in emerging markets into the mainstream. Mainstream rehabilitation services require not only advanced technology but also compelling features that will encourage therapy providers to employ them. What is compelling to therapy providers, however, may not be equally important to patients, clinical managers, clinical physicians, therapists, or payors. We therefore must pay heed to all features so that all perspectives and users will be accommodated.

In order to satisfy all perspectives and users without generating too much controversy, we believe that the success of a therapeutic neuro-rehabilitation can be defined by positive answers to all of the following benchmarks:

1. Does the therapy help?
2. Does the therapy help more than “usual” standard of care?
3. Does the therapy help more than “usual” standard of care at the same or lower cost? OR if higher cost, does it present a positive cost/benefit ratio?

Take the example of VA-ROBOTICS: the researchers compared three sets of chronic stroke patients receiving upper extremity robotic therapy, an intensity matched comparison group, and an usual care group in the VA system.

While the robotic and intensity matched therapy groups improved, the usual care group did not satisfy the first criterion: it did not lead to any measurable improvement.

The robotic and intensity matched therapy groups also satisfied the second criterion: robotic group and intensity matched group improved more than usual care.

When we benchmark these groups against each other in terms of cost, robotic therapy for the upper extremity was considerably cheaper than the intensity comparison group and it led to slightly lower overall healthcare cost (intervention plus all the healthcare utilization costs) than both the usual care and the intensity matched comparison group, thereby satisfying the third benchmark (thought not to be a significant difference).

In view of these three positive answers to our benchmarks, one can argue that upper extremity interactive robotic rehabilitation looks very promising and we suggest that the VA-ROBOTICS represents the tipping-point favoring robotic therapy and moving it into mainstream rehabilitation services.

In contrast, the same conclusion cannot be reached for the lower extremity, as evidenced by the results obtained by Hornby and Hidler when employing the Lokomat to deliver therapy for chronic and sub-acute stroke patients.

The robotic Lokomat group led to improvements in both chronic and sub-acute stroke patients satisfying the first benchmark. However, it failed to satisfy the second benchmark: compared to usual care as practiced in the US, it led to inferior results. Accordingly, it would be premature to benchmark it against the third criterion.

CONCLUSION

VA-ROBOTICS represents the tipping point of upper extremity robotic therapy. We contend that robotic therapy for the upper extremity that involves an interactive high intensity, intention-driven therapy based on motor learning principles and assist-as-needed
leads to better outcomes than usual care in chronic stroke (and probably with even a greater impact for acute/sub-acute strokes). Moreover, this treatment modality is now practical to implement in the clinical realm. But much remains to be answered and researched. We still don’t know how to tailor therapy for a particular patient’s needs. We do not know the optimal dose or in cost-benefit terms: What is the minimum intensity to promote meaningful change? Is too much therapy detrimental? Should we deliver impairment-based or functionally-based approaches and to whom: severe, moderate, mild strokes? Should we assist-as-needed, resist, or perturb and augment error? Who might be the responders who benefit most from these interventions? What adjunct therapies might have additive value? How should we integrate the robotic gyms in therapy practices?

The situation is less bright for the lower extremity. While we cannot argue with data nor with AHA or VA/DOD statements that robotic for lower extremity is still in its infancy, we believe that the field can mature and demonstrate its promise for lower extremity rehabilitation. We and other research groups are working towards such a goal of properly understanding the neuroscientific basis of gait and stroke recovery and of exploring creative solutions to move robotic therapy for the lower extremity to the same standing as upper extremity.

Acknowledgments

This work was supported in part by NIH grant #1 R01-HD045343 and the Veterans Administration Baltimore Medical Center “Center of Excellence on Task-Oriented Exercise and Robotics in Neurological Diseases” B3688R. H. I. Krebs and N. Hogan are co-inventors in several MIT-held patents for the robotic technology. They hold equity positions in Interactive Motion Technologies, Watertown, MA, USA the company that manufactures this type of technology under license to MIT.

REFERENCES


**Figure 1.**

VA-Robotics Outcomes. The left column shows the comparison of the first half of the robot training (RT) group with the usual care (UC). The right column shows the comparison of the robot and intensive comparison training (ICT) groups.

<table>
<thead>
<tr>
<th>Overall Mean Difference over 36 weeks</th>
<th>Overall Mean Difference over 36 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fugl-Meyer Robot vs. UC = <strong>2.88</strong>, p=.016</td>
<td>Robot vs. ICT = <strong>-0.58</strong>, p=.63</td>
</tr>
<tr>
<td>Wolf Robot vs. UC = <strong>-8.10</strong> sec, p=.005</td>
<td>Wolf Robot vs. ICT = <strong>2.13</strong> sec, p=.55</td>
</tr>
<tr>
<td>SIS Robot vs. UC = <strong>5.95</strong>, p=.04</td>
<td>SIS Robot vs. UC = <strong>1.19</strong>, p=.55</td>
</tr>
</tbody>
</table>
Figure 2.
Illustration of the Impact of Fixed-Model Selection on the Estimates of 36-Week Outcomes
Figure 3.
Clinical Results of Robotic Therapy in Stroke Using Lokomat. Top row shows the results with chronic stroke (enrollment > 6 months post stroke) and the bottom row shows results of subacute stroke trials (enrollment between 3 and 6 months’ post-stroke). Dark gray = conventional gait training, Light gray = Lokomat gait training. * indicates statistically significant between-group differences (p < 0.05).