Neurological assessments without clinical supervision for a rare disease

Vineet Pandey, PhD[1], Nergis C. Khan, AB[2], Anoopum S. Gupta, MD, PhD[2], Krzysztof Z. Gajos, PhD[1]
[1]Computer Science, Harvard University, Cambridge, MA, USA
[2]Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, USA

Abstract Telemedicine has become mainstream. Many telemedicine tools rely on clinician availability and expertise to perform remote consultation and diagnosis. However, the number of trained experts—like clinicians—is small; this is one of many challenges that are amplified for rare diseases compared to other disorders. To improve health equity in our societies across people with different disorders, we propose building telemedicine tools that tackle the bottleneck of limited clinician availability. Furthermore, a rare disease community might not be a homogeneous group across abilities, socio-economic status, and more. To support the range of people within a community, we suggest building processes and tools with such variation in mind. In this paper, we provide an overview of the success of the *Hevelius* system in producing high-quality clinical severity scores for a neurological rare disease community. People use the tool from home with no clinical supervision. Additionally, we provide case studies of our efforts to improve *access*, *adoption*, and *adherence* for our tool across different members of a rare disease community. With our lessons and questions, we invite feedback from the WISH community in envisioning how to design sociotechnical systems that support telemedicine health experiences without clinician involvement.

Introduction: Beyond being there in the clinic

> "The rhetoric of democratizing education implies broad social benefits without precisely articulating how those benefits might be distributed." - Hansen, Reich[1]

Many telemedicine tools support remote diagnosis with interactions between clinicians and people with disorders. *E.g.*, a neurologist might assign a motor impairment score to a person's dominant arm by looking at the video feed of the Finger-to-Nose test. Such rich synchronous interactions between clinicians and people provide promising alternatives to in-clinic consultation. However, they still rely on the critical expertise of clinicians. In the spirit of the seminal Human-Computer Interaction work *Beyond Being There*[2], we ask whether this approach of shifting in-clinic assessments to the online realm fails to consider a much broader range of possibilities for telemedicine tools? Additionally, do such online in-clinic visits improve or worsen digital health equity across different disorders? We situate our discussion in our experiences with families caring for children with a rare neurological disease while supporting remote data collection with *Hevelius* — a remote telemedicine tool that outputs valid and reliable clinical severity scores for a neurological rare disease with a few minutes of usage at home.

In-clinic visits require efforts from all, but more so from people with rare diseases

In-clinic visits have been a cornerstone of medical experience across centuries and cultures. However, the Covid-19 pandemic has pushed people to suddenly find reliance and comfort in online interactions with clinicians[3]. If not for this big shift, the act of questioning in-clinic visits might have been questionable itself. In-clinic visits require substantial efforts from people with disorders and their caregivers. These include physical efforts (such as helping the person get to the clinic), mental efforts (e.g., dealing with the stress of scheduling, planning, and executing an in-clinic visit), and financial resources (e.g., commute costs).

Many such challenges are amplified for people with *rare diseases*. And while any one rare disease affects a small number of individuals, nearly 25-30 million people in the United States are affected by some rare disease. For the neurological rare disease (Ataxia-telangiectasia, henceforth A-T) that our research team has experience with, all in-clinic visit challenges mentioned above are amplified. A-T is a progressive, life-limiting disease that manifests in childhood. Most people with A-T have motor impairments requiring assistance in movement, including the use of assistive technologies (like wheelchairs) that might require constant assistance or supervision. Scheduling in-clinic visits is made

more challenging with the limited presence (and hence, availability) of relevant clinical expertise in a geographical area[4]. Many families have reported cycling through multiple care providers to find clinicians with appropriate expertise. Finally, for many A-T families, making an in-clinic visit with a knowledgeable specialist requires traveling long distances. Furthermore, our conversations with caregivers revealed additional disruption that scheduling an in-clinic visit brought to the entire family: asking for multiple-day leave from work weeks in advance or family members taking turns seeing the doctor across different visits and creating ways to share information with each other. Finally, many people with A-T might be immunocompromised which adds extra concerns with traveling in general and more so in times of the Covid-19 pandemic. Overall, the actual time spent in an in-clinic visit is likely just a fraction to the total effort put in to making the visit possible.

The above discussion suggests that rare disease communities stand to gain a lot from telemedicine solutions. And, indeed, telemedicine encounters are likely to be more common after Covid-19 than they were before: of the 16 neurologists we surveyed recently, 13 told us that they plan to use telemedicine approaches more after Covid-19 than they did before. However, all neurologists who answered the question, told us that the exam part of a health encounter is harder to perform well in telemedicine encounters compared to in-person visits. A key reason cited was the difficulty of making an accurate and objective assessment of a person's state. Wouldn't it be great if broadly accessible tools existed to enable remote health assessments from the comfort of one's home?

System and Results

Hevelius is an online motor impairment assessment tool for people with neurological conditions. Hevelius comprises a computer mouse-based tool that provides objective, granular, interpretable, multidimensional quantification of motor impairment in the dominant arm with just a few minutes of use[5]. Compared to a standard neurological exam and existing disease rating scales, Hevelius does not require expert judgement to compute the scores and it provides assessments that are more granular. In addition to the motor impairment measurement component, Hevelius also includes questionnaires for collecting: 1) self-reports on health and lifestyle; and 2) self-reports about using the tool (Figure 1). An earlier in-clinic validation of Hevelius with people with Ataxia and Parkinsonism demonstrated that the measures produced by Hevelius can be used to accurately detect each of the two diseases and to produce accurate and granular estimates of the disease severity[5]. In this paper, we reflect on a deployment of Hevelius in the homes of families with A-T children for unsupervised use. Throughout this section and beyond, we refer to a user with neurological disorder as a participant and their family members overseeing their tool use as caregivers. We refer to a unique participant, caregiver(s) set as a family.

A 12-week deployment with 12 families with A-T children ranging in age from 6 to 16 years, demonstrated high correlation with clinical scores (Figure 2), and high test-retest reliability (not shown).

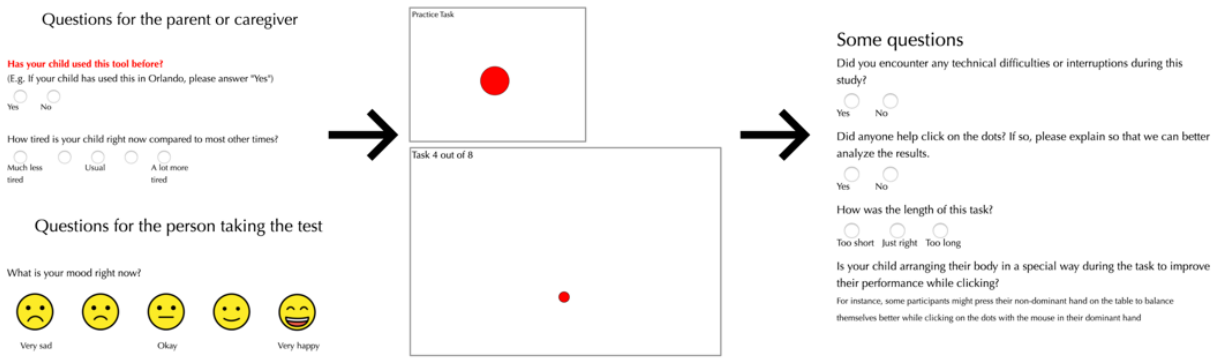

Figure 1: With Hevelius, rare disease community members provide researchers weekly well-being self-reports,motor performance data, and insights about tool usage. The user flow for Hevelius: A) Caregivers and participants answer questions about lifestyle and well-being; B) Participants perform practice tasks to warm up and then click on eight rounds of nine dot clicks; C) Caregivers answer questions about the participants'experience using the tool.

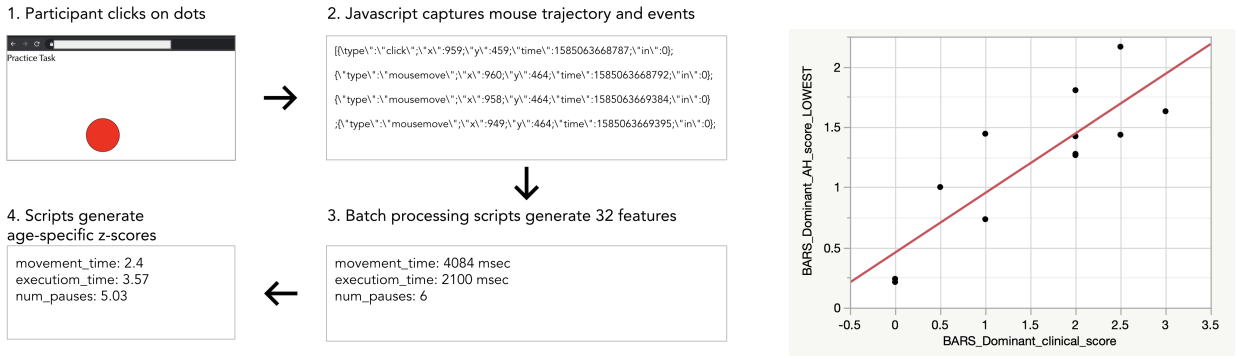

Figure 2: (a) Hevelius transforms mouse trajectory and events to 32 features and then compares these values to the normative data to provide age-specific z-scores. All these z-scores are then converted to a predicted score by a machine learning model. (b) Clinician assigned scores from an in-clinic visit correlated strongly with Hevelius prediction (based on the lowest predicted score across 7 at home sessions); r=0.85; $p<0.001$

Reflection - Does Hevelius improve/worsen health equity?

Our results from at home deployment demonstrate that people provide clinically valid motor performance assessment scores. But does the tool do so evenly for our participants? To understand this, we analyze families' self-reports and interviews according to Veinot et al.'s framework[6] and share case studies for tool access, adoption, and adherence. Conceptually, we used three levers to improve participation: 1) designing a robust tool; 2) collaborating with a rare disease foundation; and 3) providing families access to the research team for any challenges/concerns. We share how our efforts succeeded, which concerns showed up, and how do we plan to tackle them.

Access

What worked well: One family reported concerns about not being able to access the tool. The research team verified that they were using a browser with an old version; updating the browser fixed the issue. Accessing the tool requires basic (not high-bandwidth) internet connection. None of the families mentioned this as a concern. Before the study started, we provided the families with a mouse to reduce their workload and to reduce chances of not starting due to not having an appropriate device. Most people mentioned having a laptop, so that was not an issue.

Concerns: We note that the absence of evidence does not imply the evidence of absence: our participant set might have been biased. All families were recruited via the rare disease foundation that has collaborated with the research team before. It is possible that the families using the tool might not be representative of the broader A-T/US/global population; early adopters rarely are.

Steps ahead: First, We intend to understand the profiles of families that have used the tool. We propose two ways: 1) capturing demography as features (income levels, location, education attainment, occupation), and 2) create portraits of families based on their needs and their assets e.g. which support systems (if any, like grandparents) do they have at home? With this analysis, we intend to uncover demographic biases in the recruited families and understand which family strengths' can our tool play well to (and how). Second, to improve the demography of families in future deployments, we can work with rare disease foundations with a national spread; however, we currently don't know whether such foundations' access to potential participants follows a demography profile representative of people with A-T. Using nontraditional venues like libraries, faith based organizations, and community groups (as suggested in Veinot et al[6]) might not be useful in recruitment. We can likely reach out to online fora and other communities (like subreddit) with clear details and asks.

Adoption

What worked well: Participants did not report any concerns using the web interface. We provided an easy to use tool that required limited instructions. Clicking dots on a screen provides a simple, clear goal. To reduce confusion about navigation, the tool also featured a simple, linear navigation style where participants just followed the instructions or

clicked yes/no to questions. We minimized usage burden by providing large buttons (whenever not part of the severity evaluation). To make this task ability-specific, participants could select a dot size any time that they felt the current dot size was too small.

Any carefully designed tool can still face issues in new environments. To handle such cases, we provided the contact details (emails and phone number) of the lead researcher. Participants subsequently contacted the researcher to share concerns — e.g one participant likely clicked on the "view source" button (using right click, by mistake) while another found some dots to show at the edges of their small screen. In both these cases, the lead researcher debugged the issue and (for the latter) altered the tool to work across a broader set of screen sizes while not changing the usage experience or clinical validity of the tool.

Concerns: Most participants were not comfortable using the mouse. Some caregivers noted that participants needed adult supervision or encouragement to perform the task both due to the ability-specific difficulty inherent in the task and due to the frustration of using a mouse (e.g. making right clicks by mistaken when (left-)clicking on a dot). This is a factor that can alter which families continue using the tool and which ones do not. While some parents might have the time, financial, or social resources to either supervise the participant or find someone to do that, others might struggle with it. There are also emotional labor concerns from watching a loved one struggle to perform a task, further burdening some families (that might already be stressed) over others.

One participant stopped using the tool because even the largest dot size was too difficult for them to work with. This provides a limitation of work in reducing health equity within the rare disease community. Such cases are difficult to debug beforehand by simply running pilots (like HCI researchers do for many software aimed at able-bodied people). In fact, this small study is indeed a pilot for larger deployment. Rare disease participants are, by definition, difficult to find; running extensive pilots can be burdensome for them and also reduce the number of participants we might reach out. Furthermore, the heterogeneity in both the underlying condition and progression can make some severity levels difficult for the tool to work with.

Steps ahead: Should we spend more efforts making this tool work? Should we leave behind such participants?

Adherence

Our results show that people used the tool sporadically: participation dropped after seven weeks of at home usage and few families used the tool weekly as suggested.

What worked well: Before the study started, we had asked families to choose a day to use the tool every week. We felt providing families with the agency and flexibility to choose a day/time would work better than the research team fixing such details. Two members of the research team discussed tool usage data every Monday and noted which participants had not used the tool in a while (roughly 3 weeks). To improve adherence, we shared the participants' code with the rare disease foundation who sent out emails to remind the families about using the tool and to understand if they were facing any challenges our team could assist with. Initially, we saw an increase in tool usage and heard back from participants when the emails went out. However, both the usage and participant reports dropped over time.

Concerns: As a research prototype, the tool is available free of cost. But do people have the "free" time to do this? We might have misjudged the broad range of time-intensive responsibilities that caregivers perform for family members with and without A-T, themselves, and their professional work. Most participants reported lifestyle factors—e.g. travel, busyness—for not adhering every week. Furthermore, the most appropriate time in a day to use the tool (when the participant felt alert) might not have correlated well with the time when caregivers have more leisure time.

Steps ahead: Kid-friendly games within the tool can motivate tool usage and reduce the need for adult supervision; we are building such games now.

Conclusion

In this paper, we discussed the Hevelius tool and our processes towards supporting remote health assessments for a neurological rare disorder. As we consider scaling our tool's deployment to answer more clinically and socially relevant research questions, we look forward to the WISH community's critique of our work.

Acknowledgements

This work was supported in part by the Ataxia-Telangiectasia Children's Project and by NIH grant 1R01CA204585-01 as part of the NSF/NIH Smart and Connected Health program. We thank Winnie from Massachusetts General Hospital, and Sarah and Jennifer from A-T Children's Project for their invaluable help in communicating with the A-T community members.

References

1. Hansen JD, Reich J. Democratizing education? Examining access and usage patterns in massive open online courses. Science. 2015 Dec 4;350(6265):1245-8.
2. Hollan J, Stornetta S. Beyond being there. In Proceedings of the SIGCHI conference on Human factors in computing systems 1992 Jun 1 (pp. 119-125).
3. Hollander JE, Carr BG. Virtually perfect? Telemedicine for COVID-19. New England Journal of Medicine. 2020 Apr 30;382(18):1679-81.
4. Jacobs M, Gheihman G, Gajos KZ, Gupta AS. " I think we know more than our doctors" How Primary Caregivers Manage Care Teams with Limited Disease-related Expertise. Proceedings of the ACM on Human-Computer Interaction. 2019 Nov 7;3(CSCW):1-22.
5. Gajos KZ, Reinecke K, Donovan M, Stephen CD, Hung AY, Schmahmann JD, Gupta AS. Computer mouse use captures ataxia and parkinsonism, enabling accurate measurement and detection. Movement Disorders. 2020 Feb;35(2):354-8.
6. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. Journal of the American Medical Informatics Association. 2018 Aug;25(8):1080-8.
7. Wong-Villacres M, DiSalvo C, Kumar N, DiSalvo B. Culture in Action: Unpacking Capacities to Inform Assets-Based Design. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems 2020 Apr 21 (pp. 1-14).