JMIR Biomedical Engineering

Engineering for health technologies, medical devices, and innovative medical treatments and procedures
Volume 7 (2022), Issue 1 ISSN 2561-3278 Editor in Chief: Syed A. A. Rizvi, MD, PhD, MBA, MPH, BSN

Contents

Viewpoints

Original Papers

Viewpoint

Reducing Treatment Burden Among People With Chronic Conditions Using Machine Learning: Viewpoint

Harpreet Nagra[1*], PhD; Aradhana Goel[2*], MS; Dan Goldner[1*], MEd, PhD

[1]One Drop, New York, NY, United States

[2]Integrated Care, Bayer Pharmaceuticals, San Francisco, CA, United States

[*]all authors contributed equally

Corresponding Author:
Dan Goldner, MEd, PhD
One Drop
166 Mercer St, 2nd Floor
New York, NY, 10012
United States
Phone: 1 800 437 1474
Email: dan@onedrop.today

Abstract

The COVID-19 pandemic has illuminated multiple challenges within the health care system and is unique to those living with chronic conditions. Recent advances in digital health technologies (eHealth) present opportunities to improve quality of care, self-management, and decision-making support to reduce treatment burden and the risk of chronic condition management burnout. There are limited available eHealth models that can adequately describe how this can be carried out. In this paper, we define treatment burden and the related risk of affective burnout; assess how an eHealth enhanced Chronic Care Model can help prioritize digital health solutions; and describe an emerging machine learning model as one example aimed to alleviate treatment burden and burnout risk. We propose that eHealth-driven machine learning models can be a disruptive change to optimally support persons living with chronic conditions.

(JMIR Biomed Eng 2022;7(1):e29499) doi:10.2196/29499

KEYWORDS

artificial intelligence; machine learning; behavioral science; chronic conditions; self-care; behavior; chronic disease; prediction; algorithm; lifestyle; adapt

Introduction

The COVID-19 pandemic has surfaced multiple concerns present within our health care systems, including the high infection risk prevalent among people with chronic conditions, and the fact that practitioners can only provide specialized responses to acute illnesses [1]. These, in turn, leave people with chronic conditions to experience fragmented, poorly coordinated, and limited support in their treatment [2], which exacerbates the treatment burden patients experience as they encounter decreased support for their ongoing medical care [3]. Increased treatment burden can heighten the risk for illness-related burnout—a chronic affective state comprising symptoms of emotional exhaustion, physical fatigue, and cognitive weariness, often the outcome of depletion of energetic resources resulting from prolonged exposure to medical distress [4]. Considered the most efficacious of the various chronic illness frameworks [5], the Chronic Care Model (CCM) [6,7] addresses how health care teams, including physicians, can better support those with chronic conditions by shifting care focus to coordinated self-management and decision-making support [8-10].

eHealth supports individuals in self-care and facilitates interactions and collaboration within the health care system, thereby reinforcing the value of the CCM. eHealth technologies are connective elements that build bridges between stakeholders in the ecosystem [11]. The eHealth enhanced Chronic Care Model (eCCM), developed in 2015 by Gee et al [12], is a framework that incorporates eHealth literature into the CCM components and promotes understanding how eHealth tools such as mobile health (mHealth) apps, machine learning (ML), e-communities, electronic health records, and eHealth education may facilitate the implementation of the CCM in a digital space (eg, by enabling self-tracking of health data, empowering involvement in shared decision making, supporting preparation for appointments, and enabling personalized decision support

through visualization of data and reminders). Thus, there are various technologies that have been suggested to improve the functionality of the different components of the CCM.

Of the suggested technologies, ML offers new opportunities to deliver more accessible, equitable, personalized, and cost-efficient chronic care programs. ML may help mitigate treatment burden and burnout risk by providing self-management and decision-making interventions that guide and support people with chronic conditions. This guidance and support can be delivered directly or via a mobile app; it can also be provided to a coach or care provider to consider when working with a person with chronic conditions.

Here, we introduce an emerging ML model called "the outcomes model" as an example of the eCCM framework, which integrates behavior and data science principles to reduce treatment burden and chronic-condition–related burnout. The outcomes model correlates an individual's health outcomes months in advance with current lifestyle and biometric markers and could be used to help determine which kinds of lifestyle activities (eg, maintaining a sleep routine or adding more daily physical activity) are most likely to benefit an individual at a given time. It may not only help people make meaningful lifestyle decisions but may also enhance their motivation through self-monitoring and feedback on health behaviors and outcomes in order to complete activities necessary to improve chronic-condition–related self-management.

Machine Learning at One Drop

One Drop is a mobile app that passively and actively collects data (passively from wearables such as Apple Watch and FitBit, and actively from user-inputted information directly into the mobile app). It provides ML health trend prediction, tailored educational content, and personalized health coaching services for people who want to prevent and manage chronic conditions, such as diabetes and heart disease. Upon downloading the app, members actively grant permission to use the data only for app improvements and for research and development of ML models for use within the app.

One Drop data collection spans 195 countries, and as of this writing, the data comprise over 30 billion data points. Some of the categories and their components are presented in the following:

1. Biometrics data: heart rate, blood pressure, skin temperature, and blood glucose concentration.
2. Behavior data: physical activity, sleep, food and mealtimes, medications and consistency of medication use, and geographic movement.
3. Outcome data: laboratory measures, such as glycosylated hemoglobin type A1c (HbA1c) and cholesterol levels, and self-reported metrics, such as weight, body mass index, and waist circumference.
4. Engagement data: app use, interactions with coaches, use of educational resources, interactions with peer support networks, response rate to app notifications, and others.

Information is collected on multiple cadences; some wearable data are collected passively, while others are set by the person logging manually. App features call for medical-condition–specific subsampling strategies. More frequent data input supports more accurate outcomes forecasts.

Currently, ML algorithms trained on these data predict health outcomes, such as weight, average blood glucose concentration (HbA1c), time-in-range, or blood pressure, using the inputs described above (eg, biometrics, behavior, outcomes, and engagement), for 1-6 months in advance. One Drop data from a sample of over 50,000 app users were used to train a suite of patent pending supervised learning models, each for a different health metric (weight, blood pressure, or blood glucose) and time horizon (1-2 months, 2-3 months, 3-4 months, or 4-6 months). Data collected prior to 2019 were used to train the algorithms; data from January 2019 through February 2020 were used for testing. Test set predictions were 10%-40% more accurate than a naïve (benchmark) prediction of "no change." Further details of that work have been presented previously [13,14], and model development is still in progress. For the present discussion, we note this as an example that informative biometric predictions based in part on behavioral inputs exist today. However, predictions such as these cannot in and of themselves produce improved outcomes nor reduce treatment burden. Only in combination with a behavioral support framework can such benefits be realized. We next describe such a combination, which we call the outcomes model.

The Outcomes Model

For people with chronic conditions, medical appointments with health care providers may occur on a quarterly or biannual basis. Having a system such as the outcomes model could provide lifestyle support in between medical visits, while simultaneously giving people with chronic conditions a predictive insight into how their habits are impacting their overall health. To date, no eHealth offerings focused on people with chronic conditions integrate a combination of multiple ML models to (1) forecast outcomes up to 6 months into the future; (2) provide insight into which lifestyle behavior modification may most significantly impact a person's desired clinical outcome; and (3) use various data inputs (eg, biometrics, behavior, outcomes, and engagement) to support behavior change and reduce treatment burden for people with chronic conditions.

The hypothesis we are exploring is how to use an ML model capable of predicting likely changes in outcomes to reduce treatment burden within the eCCM. A potential approach is to use a forecasting model to determine which lifestyle modifications are most likely to yield the greatest improvement on forecasted clinical outcomes of interest (eg, weight, HbA1c, blood pressure, and time-in-range). Based on the forecast of various health outcomes-focused interventions, an optimal current lifestyle modification focus could be selected. A combination of human and automated interventions could then be initiated to suggest adjustments to the individual's behavior. After an initial trial period to evaluate the effect of a suggestion, the guidance could be recalculated with up-to-date information, and the focus might either be maintained or switched to a new, now most optimal, choice. If the recommended focus is not practical to be addressed by people with chronic conditions for

any reason, the next most effective mode could be selected. In this way, the system can support people with chronic conditions in making progress toward better health outcomes using tailored interventions that adapt to each individual, evolve with time, and are informed by their predicted effect on the individual's health.

Predictive guidance, developed as described above, might be digitally delivered directly to people with chronic conditions, which can support the calculated guidance provided through a healthcare professional. This predictive guidance could reduce treatment burden by helping people with chronic conditions seek additional help from a health care professional. The information from predictive guidance could be used more broadly as a part of a health care network's decision support system. Prescribers would be required to evaluate potential medical interventions such as medication dose changes. Lifestyle modifications, such as modifying physical activity or carbohydrate intake, could be improved upon by a digital health coach, a dietitian, or other subject matter experts. These coaches are typically more readily available for appointments between clinic visits, and the use of their services may reduce the burden on the health care system by increasing patient engagement in self-care and supporting health care providers with additional patient health insight.

Discussion

Machine Learning-Based Lifestyle Modifications

The general idea of informing chronic condition management with computed forecasts about an individual is receiving increasing attention. Schwartz et al [15] review a variety of current developments based on the rise in quantification of many aspects of daily health. One of the many features they see as important in the evolution of digital health is prioritization of interventions, and in that light, they review concepts such as ecological momentary interventions that "assess the person's (digital twin) status and the model delivers interventions as needed, when needed—perhaps even preemptively." More recently, Chevance et al [16] have elaborated on emerging applications that use predictions to support behavior change, with examples including predictions of smoking or walking behaviors. One aspect of their discussion centers on continuous tuning interventions, that "include real-time optimization algorithms, which [based on such forecasts] can further adjust intervention content or delivery aspects to the needs of a specific individual." Increased attention notwithstanding, we concur with the assessment by Chevance et al that "Computational models and associated computerized simulations are still relatively under-used in the behavior change field."

A wide set of ideas is discussed in those papers, and we are broadly supportive of all of them. The approach we have described here sits squarely among those ideas. While Schwartz et al [15] describe the role of algorithms for informing a digital phenotype, their digital phenotype is not necessarily predictive of a future state. Chevance et al [16] flesh out continuous tuning based on updated predictions, but their examples focus on behavioral interventions tuned via behavioral predictions. Our aim here is to advocate specifically for placing biometric predictions in a central role in steering behavioral interventions. This stance is tantamount to a hypothesis that a dominant contributor to treatment burden for people with chronic conditions is uncertainty. Will a particular behavior change succeed in improving my condition, given my history and circumstances? Is the work I am doing now truly going to make a difference later? We believe that definitive answers to those personal questions, updated as individual behaviors, the chronic condition and circumstances evolve, can make a substantial reduction in treatment burden. We further believe that predicting the biometric effect of candidate behavioral interventions to each specific individual is the most promising way to obtain those answers.

Here, we have described a strategy in which a biometric forecasting model is used to select the lifestyle intervention most likely (at a given moment) to improve an individual's forecast, and to use that as the basis for informing interventions. We acknowledge that this strategy is not a straightforward exercise. People with chronic conditions do not typically have the privilege to make one behavior change at a time. They must instead focus on making multiple changes (eg, diet, physical activity, and medication) simultaneously. Type 1 diabetes provides a good case in point for this issue. If a person living with type 1 diabetes decides to engage in more physical activity, they must make simultaneous changes to carbohydrate intake, as well as adjusting insulin-to-carbohydrate ratios. In this scenario, a suggested ML model approach is to observe the population of mobile app users for typical correlations with changes to behavior; the next step is to apply correlated factors to hypothetical changes of the past 30 days of an individual's observed behavior and examine if there are any resulting changes to the forecasted outcome. It is, to be sure, more difficult to learn from data where many recorded behaviors are correlated and overlapping, calling for careful statistical evaluation of whether the resulting forecast changes differ significantly. In these circumstances, the results and recommended behaviors can be less specific and more complex—less than ideal if the goal is to reduce treatment burden and burnout by making clearer recommendations that are more likely to succeed. Nonetheless, whether choosing a single behavior change or a combination, knowing which choice is forecasted to result in greater improvement for a given individual could inform both guidance and goal setting for that individual. Such a result would not constitute proof or a guaranteed outcome. However, the vast data reservoir on which the ML model is trained offers the possibility of learning which interventions have led to greater improvements among people in similar circumstances to the individual in question.

While recognizing the challenges, we believe that relieving treatment burden cannot improve beyond a limited threshold without reducing uncertainty about the effectiveness of a behavioral intervention for each individual at each moment in time. Fusing biometric prediction with behavioral science frameworks such as eCCM is, in our view, the best strategy for reducing that uncertainty. Whether through a predictive guidance system such as the outcomes model described above or in some other form, adjustments to behavioral interventions must be informed by an individualized prediction of which adjustments

are most likely to succeed. Adjustments so informed are necessary for sustainable chronic illness prevention, management, and hopefully treatment burden alleviation.

Conclusion

ML-based biometric predictions used in the context of established behavior change frameworks offer exciting potential to support and reduce treatment burden, as well as mitigate burnout risk for those living with chronic conditions. Chronic care management requires constant attention, which necessitates deep engagement with supportive tools. eHealth solutions such as the outcomes model may break down the boundaries that define traditional, nondigital care. Such innovation should support digital health care's progression out of reactive and into proactive chronic care.

Conflicts of Interest

AG is employed by Bayer Pharmaceuticals. HN and DG are employed by One Drop.

References

1. Sisó-Almirall A, Kostov B, Sánchez E, Benavent-Àreu J, Paz LG. Impact of the COVID-19 Pandemic on Primary Health Care Disease Incidence Rates: 2017 to 2020. Ann Fam Med 2022 Sep 24;20(1):63-68 [FREE Full text] [doi: 10.1370/afm.2731] [Medline: 34561213]
2. Wagner EH. Organizing Care for Patients With Chronic Illness Revisited. Milbank Q 2019 Sep 19;97(3):659-664 [FREE Full text] [doi: 10.1111/1468-0009.12416] [Medline: 31424130]
3. Sav A, McMillan SS, Kelly F, Kendall E, Whitty JA, King MA, et al. Treatment burden among people with chronic illness: what are consumer health organizations saying? Chronic Illn 2013 Sep 23;9(3):220-232. [doi: 10.1177/1742395312463411] [Medline: 23093542]
4. Armon G, Melamed S, Toker S, Berliner S, Shapira I. Joint effect of chronic medical illness and burnout on depressive symptoms among employed adults. Health Psychol 2014 Mar;33(3):264-272. [doi: 10.1037/a0033712] [Medline: 23895204]
5. Davy C, Bleasel J, Liu H, Tchan M, Ponniah S, Brown A. Effectiveness of chronic care models: opportunities for improving healthcare practice and health outcomes: a systematic review. BMC Health Serv Res 2015 May 10;15(1):194 [FREE Full text] [doi: 10.1186/s12913-015-0854-8] [Medline: 25958128]
6. Glasgow RE, Orleans CT, Wagner EH. Does the chronic care model serve also as a template for improving prevention? Milbank Q 2001 Jun 06;79(4):579-612, iv-v [FREE Full text] [doi: 10.1111/1468-0009.00222] [Medline: 11789118]
7. Bodenheimer T. Interventions to improve chronic illness care: evaluating their effectiveness. Dis Manag 2003 Jun;6(2):63-71. [doi: 10.1089/109350703321908441] [Medline: 14577900]
8. DiPiero A, Dorr DA, Kelso C, Bowen JL. Integrating systematic chronic care for diabetes into an academic general internal medicine resident-faculty practice. J Gen Intern Med 2008 Nov 28;23(11):1749-1756 [FREE Full text] [doi: 10.1007/s11606-008-0751-5] [Medline: 18752028]
9. Dorr D, Wilcox A, Donnelly S, Burns L, Clayton P. Impact of generalist care managers on patients with diabetes. Health Serv Res 2005 Oct;40(5 Pt 1):1400-1421 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00423.x] [Medline: 16174140]
10. Musacchio N, Lovagnini Scher A, Giancaterini A, Pessina L, Salis G, Schivalocchi F, et al. Impact of a chronic care model based on patient empowerment on the management of Type 2 diabetes: effects of the SINERGIA programme. Diabet Med 2011 Jun;28(6):724-730. [doi: 10.1111/j.1464-5491.2011.03253.x] [Medline: 21294769]
11. Botti A, Monda A. Sustainable Value Co-Creation and Digital Health: The Case of Trentino eHealth Ecosystem. Sustainability 2020 Jun 29;12(13):5263. [doi: 10.3390/su12135263]
12. Gee PM, Greenwood DA, Paterniti DA, Ward D, Miller LMS. The eHealth Enhanced Chronic Care Model: a theory derivation approach. J Med Internet Res 2015 Apr 01;17(4):e86 [FREE Full text] [doi: 10.2196/jmir.4067] [Medline: 25842005]
13. Wexler Y, Goldner D, Merchant G, Hirsh A, Huddleston B, Dachis J. One- to Six-Month Outcomes Forecasts for Diabetes and Related Conditions. American Diabetes Association 80th Scientific Sessions. Arlington, VA: American Diabetes Association; 2020 Jun 13. URL: https://plan.core-apps.com/tristar_ada20/abstract/7b03f725-d051-4ca7-8b00-084c4824bbd6 [accessed 2021-12-10]
14. Wexler Y. One- to Six-Month Forecasts of Time-in- Range: Poster Presentation. Diabetes Technology Meeting Abstracts. 2020 Nov. URL: https://journals.sagepub.com/doi/10.1177/1932296821996093?icid=int.sj-abstract.similar-articles.2 [accessed 2022-02-04]
15. Schwartz S, Wildenhaus K, Bucher A, Byrd B. Digital Twins and the Emerging Science of Self: Implications for Digital Health Experience Design and "Small" Data. Front. Comput. Sci 2020 Oct 15;2:31. [doi: 10.3389/fcomp.2020.00031]
16. Chevance G, Perski O, Hekler E. Innovative methods for observing and changing complex health behaviors: four propositions. Transl Behav Med 2021 Mar 16;11(2):676-685 [FREE Full text] [doi: 10.1093/tbm/ibaa026] [Medline: 32421196]

Abbreviations

CCM: Chronic Care Model

eCCM: eHealth enhanced Chronic Care Model
HbA1c: glycosylated hemoglobin, type A1C
mHealth: mobile health
ML: machine learning

Edited by A Mavragani; submitted 09.04.21; peer-reviewed by Y Du, L Guo, C Stecher; comments to author 29.04.21; revised version received 01.10.21; accepted 20.01.22; published 10.02.22.

Please cite as:
Nagra H, Goel A, Goldner D
Reducing Treatment Burden Among People With Chronic Conditions Using Machine Learning: Viewpoint
JMIR Biomed Eng 2022;7(1):e29499
URL: https://biomedeng.jmir.org/2022/1/e29499
doi:10.2196/29499
PMID:38875589

Viewpoint

Democratizing Global Health Care Through Scalable Emergent (Beyond the Mobile) Wireless Technologies

Graham B Jones[1], PhD, DSC; Andrew Bryant[2], PhD; Justin Wright[3], PhD

[1]Technical Research and Development, Novartis Pharmaceuticals, East Hanover, NJ, United States

[2]Technical Research and Development, Novartis AG, Basel, Switzerland

[3]Global Drug Development Connected Health, Novartis Pharmaceuticals, East Hanover, NJ, United States

Corresponding Author:
Graham B Jones, PhD, DSC
Technical Research and Development
Novartis Pharmaceuticals
1 Health Plaza
East Hanover, NJ, 07936
United States
Phone: 1 8572757045
Email: graham.jones@novartis.com

Abstract

Advances in mobile phone technologies coupled with the availability of modern wireless networks are beginning to have a marked impact on digital health through the growing array of apps and connected devices. That said, limited deployment outside of developed nations will require additional approaches to collectively reach the 8 billion people on earth. Another consideration for development of digital health centered around mobile devices lies in the need for pairing steps, firmware updates, and a variety of user inputs, which can increase friction for the patient. An alternate, so-called Beyond the Mobile approach where medicaments, devices, and health services communicate directly to the cloud offers an attractive means to expand and fully realize our connected health utopia. In addition to offering highly personalized experiences, such approaches could address cost, security, and convenience concerns associated with smartphone-based systems, translating to improved engagement and adherence rates among patients. Furthermore, connecting these Internet of Medical Things instruments through next-generation networks offers the potential to reach patients with acute needs in nonurban regions of developing nations. Herein, we outline how deployment of Beyond the Mobile technologies through low-power wide-area networks could offer a scalable means to democratize digital health and contribute to improved patient outcomes globally.

(JMIR Biomed Eng 2022;7(1):e31079) doi:10.2196/31079

KEYWORDS

LPWAN; wireless communication; global health; mobile technology; mHealth; wireless; cloud-based; personalization; cost; security; convenience; IoT; Internet of Things; global health

Introduction

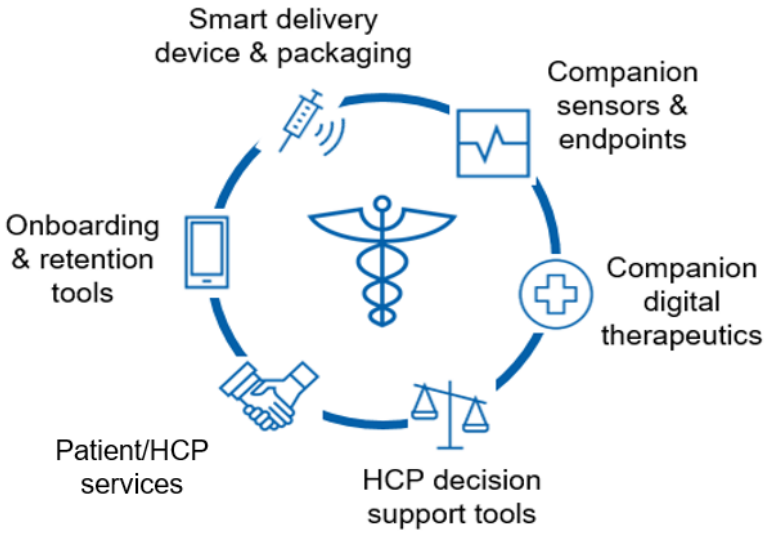
The 1962 song entitled *Return to Sender* by Elvis Presley harkens back to an age when people communicated through letter writing, and the postal service was the principal artery of our data distribution network [1]. Lest we forget, there are several features of that communication method that are noteworthy and can still inspire innovation within our modern world of telecommunications. First, the cost of the communication was borne upfront by the originator, by using either a stamp or franking imprint. The recipient's address was unambiguous, and the act of opening the communication was personal and often ethereal. There is also the option to respond to the sender at no cost with a return letter through an enclosed "self-stamped" return envelope. The modern equivalents are of course electronic, offering near-unlimited speed, scale, and customization, though this comes at a cost. The first and most obvious is a lack of personal touch. Although this may have little consequence in most circumstances, in the case of health care, we must be guided by the wishes, needs, and proclivities of a potentially unwell person whose perspectives may differ from those of a well-intentioned app developer. Second, modern communications rely on mobile devices that can be expensive (capital expenditure), have limited shelf life (depreciation), and are coupled through networks that may require monthly subscriptions (operational expenditure). In addition to these financial burdens, the devices rely on availability of increasingly

complex wired and wireless networks that have colored our expectations dramatically. Whereas a delay of one day for receipt of a letter might have been a source of irritation in days gone by, a lack of internet connectivity for mere minutes now has the potential to drive consumers into a state of frenzied panic. More problematically, the market push for ever-increasing device capabilities and networks to support them has implications for global health. There is clear evidence of a digital divide emerging between affluent and developing nations, yet the push toward digital health has the potential to impact the health care of some of the most vulnerable citizens on the planet [2]. We can all recall the frustrations of having to upgrade operating systems on our personal computers to keep pace with the visions and edicts of software developers. In the case of health care, however, such scenarios are less tenable, requiring us to plan carefully and thoughtfully how we embrace this opportunity. We advocate that the correct combination of digital services coupled with personal actions and experiences may be the solution to this conundrum and outline a case herein.

Beyond the Mobile

The promise of digitally enabled health care is being realized at an incredible pace [3]. Advances in device and broadband network technology are revolutionizing how we capture, store, and access health care data. The range of possibilities is ever expanding, impacting patients, health care providers, and payers with the establishment of Connected Health ecosystems (Figure 1). Consumer interest has been accelerated by the availability of apps on smartphones, allowing patients to access and visualize health-related data and make decisions on aspects of their health care [4].

Figure 1. Functional components of a modern Connected Health ecosystem. HCP: health care provider.

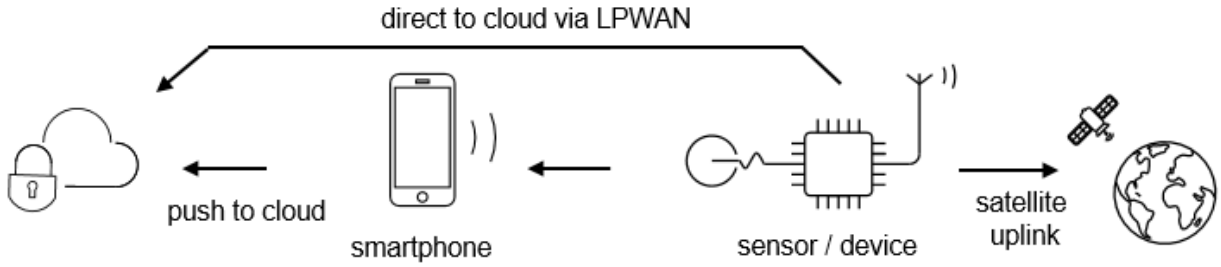


It is estimated there are approximately 3 billion smartphone users globally and over 318,000 health-related apps available for personal use [5]. That said, the highest concentration is found in developed, affluent nations and their urbanized regions, reflecting demographic trends and mirroring geographic access to quality health care services [6]. This has even led to categorization of a subclass of “fit-rich” individuals who also possess smart monitoring devices (eg, wrist-worn wearables), have ready access to a spectrum of preventative through palliative health care services, and actively engage in digital health programs [7]. Despite the seemingly limitless opportunities, the impact on health outcomes at the population level has been smaller than initially anticipated and device-related limitations have been cited as a likely culprit. Several surveys underscore patients’ diminished enthusiasm for devices over time, including app burnout, with many using features less than 3 times [8]. This presents a conundrum for the connected health community at large. Given the smartphone-centric culture of Generation X, Millennials, and Generation Z, it is logical to assume that connected health products represent an enormous opportunity, and this will likely be the case for a large segment of the health care system. However, at the population level, there are a number of inherent hurdles and limitations that need to be considered a priori for patients to use the smartphone as the central hub in their health care nexus. Cost is a factor, with premium devices exceeding US $1000, and data plans of up to US $100/month. In the United States, the average user changes their smartphone approximately every 24 months, often amortizing sunk cost into a newer model [9]. Among limitations of the base unit, battery life is often a limiting factor, with devices typically requiring daily charging despite continual advances in lithium-ion technologies. This problem is even worse for some wrist-worn systems (eg, the Apple Watch), as their need for regular recharging has been a limiting factor in clinical trials, as on-body data cannot be gathered while the device is charging [10]. The situation becomes even more complex when the smartphone is paired with additional devices—for example, when a smartphone is communicating with a drug substance supplied in a smart package, allowing patients to obtain instructions for use via wireless (near-field or Bluetooth) communication. Another example is a connected drug delivery device such as a smart autoinjector or inhaler. In both cases, in addition to the installation of a customized app, a pairing process needs to be conducted, which can be a source of friction, and requires the smartphone to be in close proximity, with sufficient battery life and wireless connectivity. Longer term, the need for firmware and app updates is likely, requiring a degree of technical

knowledge, and there may be potential interoperability limitations if contemplating switching operating systems (eg, from iOS to Android). On top of this, omnipresent security concerns arise, both for the operation of the device and the data it collects then distributes. Though the growth and use of health-related apps on smartphones seems likely to continue unabated, there are merits to using alternate and transparent approaches to connect devices, which has led to the emergence of Beyond the Mobile (BTM) as a viable proposition [11].

In this scenario, a sensor on a wireless device communicates directly to the cloud through a low-power wide area network (LPWAN) or through a satellite uplink without the need for the smartphone-based intermediate step (Figure 2). As just a few examples, a smart pill dispenser could confirm when a patient administers a drug through activation of a container cap sensor, a motion sensor might alert a caregiver to a patient's sudden fall, or a wrist-worn cardiac monitor might signal an irregular event in need of further scrutiny (eg, arrythmia). Though it is equally viable to capture this information via a smartphone, there are several advantages to a BTM approach. For example, the sensing device could be configured to work "out of the box" with minimal setup and no requirement for a separate device (and thus no need for downloading apps and updates, pairing, etc). In addition, it could be personalized to a user profile up front prior to shipping. Data could be transmitted near instantly via an appropriate network (vide infra) at no cost to the patient, with lessened security concerns as data are not stored on an uncontrolled system, nor transmitted over the open internet.

Figure 2. Uploading data from IoMT devices directly or via smartphone. IoMT: Internet of Medical Things; LPWAN: low-power wide-area networks.



Additionally, validation of such a device would be less demanding than a smartphone-paired version, as the likelihood of errors from software of unknown provenance is reduced. If debugging was proven necessary, those updates would be confined to the device itself and would not be compounded by issues relating to the smartphone and its various systems. Clearly, there are myriad benefits to such an approach, and the impact of reduced friction for the patient is noteworthy. It is well recognized that to evaluate the true impact of connected health on real-world outcomes, acquiring quality data longitudinally is key, and low friction/passive monitoring approaches are beneficial [12]. This in part has led to the development of wide arrays of newer Internet of Medical Things (IoMT) devices (eg, voice assistants), which can gather data with minimal user inputs [13].

Using BTM approaches to connected care has even deeper ramifications when considering population health at a global scale. Leaving aside limitations of smartphones themselves, their use depends on the availability of cellular networks on which the devices operate. It is estimated that nearly half of the world's population does not have access to the internet and an even greater percentage does not have access to broadband (cellular or Global System for Mobile communication) networks [14]. The problem is exacerbated in developing nations—whereas some 73% of North Americans have mobile connectivity, the lower proportions of mobile connectivity in Latin America (50%), the Middle East and North Africa (38%), South Asia (28%), and sub-Saharan Africa (21%) highlight the gaps that exist [15]. Although the deployment of 3G, 4G, and now 5G networks continues apace, the economic considerations involved mean that many areas, particularly rural, will remain cold spots for the foreseeable future. In a move to address this infrastructure gap, low-energy long-range wireless technologies are being developed and evaluated, leading to LPWANs [16]. Coupling to these networks through low-energy devices offers a very real and scalable option for impacting connected health in developing nations.

Long-range Wide Area Networks

First developed by Cycleo in France, LoRa (from "long range") represents a long-range, low-power wireless protocol that can be deployed through LPWANs [17]. Subsequently acquired by Semtech and incorporated into a global nonprofit, LoRa Alliance, the system employs sub-gigahertz radio frequency bands (principally 433-923 MHz) to allow for long-range data transmission (>15 km rurally) using low power and transfer rates of up to 30 kbit/s [17]. Long-range wide area networks (LoRaWAN) have multiple attributes, which include the following: (1) multikilometer range capacity, (2) low power use through micro batteries, (3) low entry and operating costs (capital expenditure approximately US $20, operational expenditure negligible), (4) low bandwidth requirements (250 bit/s to 11 kbit/s), (5) security available through 128-bit encryption, and (6) near limitless geographic coverage of gateways. There are limitations including low uplink/downlink speeds, favoring methods that parse data into appropriate packets for transmission with intervals, as well as edge computing methods to preanalyze data. Perhaps more importantly in this context, however, the technology uses unlicensed radio bands, and the network technology is relatively simple and affordable to access. The low cost of implementation and lack of formal network licensing leads to the potential for community-driven solutions.

New chipsets are continually evolving, with ever increasing transmission capacity coupled with decreasing size and power

consumption which, when augmented with geolocation capability, has resulted in LoRaWAN fast becoming a connection method of interest for Internet of Things devices in remote areas. Following development in Europe, LoRaWAN has been deployed in over 100 countries, which includes many in Asia, Africa, and Latin America [17]. One of the most ambitious deployments to date was undertaken by Tata Communications in India, with a goal to reach 400 million people in 2000 communities [18]. Based on promising data, numerous developing nations are now in the process of assessing how to integrate autonomous LoRaWAN base stations alongside Wi-Fi community networks to allow seamless data transfer and access [16]. In the United States, Amazon is deploying LoRa for its Sidewalk networking platform, which has been likened to a crowdsourced wireless network for neighborhoods [19]. Semtech has also launched its YoSmart and YoLink platforms to allow home connectivity using LoRa-based smart hubs, with up to a quarter-mile radius [20]. Collectively, the smart home market is expected to reach nearly US $155 billion by 2023, with the United States accounting for 40% of this, ensuring a rich innovation pipeline for subsequent global adoption [21]. Given that health care costs represent considerable proportions of gross domestic product (United States 16.9%, Switzerland 12%, sub-Saharan Africa 6.1% [ranging from 2.14%-16.1%], China 5%, India 3.6%, Organisation for Economic Co-operation and Development [OECD] average 8.8%), there is natural interest in deployment of these technologies as a component of digital health [22].

IoMT Devices

Growth in wireless and mobile technologies has been particularly strong in the health sector, establishing a new category denoted the IoMT [23]. By current estimates, there are more than 3.7 million IoMT devices in use currently and a report by Allied Market Research forecasts IoMT within the health care market will grow to US $332 billion by 2027 [24]. This spans a range of handheld and body-worn devices (smart pad, smartphone, wrist-worn trackers, smart rings), wireless sleep monitors, smart refrigerators (which monitor dietary intake), and voice analyzers capable of diagnostic interrogation (eg, Alexa, Siri). It is easy to see how communication with residential IoMT devices can fuel digital health applications given that many individuals spend the majority of the day at home (and often almost the entire day during the COVID-19 pandemic). Rich composite data streams are available and could be useful for addressing disease-specific parameters (eg, diet and exercise activity for cardiometabolic illness, or voice intonation and social network activity in depressive disorders) and also for providing interventions (eg, reordering prescription medications, voice- or video-guided instructions for use of medicaments). Driving this growth are new technologies that are ever more powerful, miniaturized, and affordable. For example, tags (using near-field communication or radio frequency identification technology) as small as 1 mm^3 can now be molded into drug packaging during manufacture, allowing patients to track delivery and access information from their smartphone [25]. Bluetooth modules mere millimeters square with shelf lives of many years can be added to devices and activated by a user—for example, on a smart pack to confirm a drug dose has been administered [26]. The deployment of these and "ready to connect" devices that have eSIMS integrated for over-air activation promises to greatly aid our efforts to monitor the health care supply chain [27], counterfeiting and security, and drug and diagnostics adherence rates. Critically for individuals in developing nations, many of these options do not require the use of a mobile phone. Direct-to-cloud capabilities are available in many areas, and combinations of fog and edge computing (allowing local processing of data prior to pushing to the cloud) are available in others [28]. Additionally, voice messaging from devices can be activated by the recipient of a package, requiring only simple microprocessors coupled to a piezo speaker and powered by a button cell. There even exists the option to record a simple return message onto such a device, which would then be transmitted to a base station (eg, via a LoRa network in small data packets). Such "return to sender" capabilities in essence close the loop between the patient and health care provider, confirming the participatory component of P4 medicine (predictive, preventive, personalized, and participatory) [29]. Ultimately, outcome measures will dictate the degree to which such connectivity will be adopted. However, early research on adherence rate improvements among patients using smart connected autoinjectors bodes well for deployment [30].

Use Cases of BTM

Myriad exciting opportunities present themselves for BTM technology to enhance patient care and advance managed health care at the community level. From the drug product side, the ability to track supplies down to the individual package level will allow deployment of state-of-the-art logistics approaches developed in the online retail industry. Receipt by the patient at home, at a pharmacy, or care facility is possible, independent of how it was transported (truck, mail, parcel pod, drone, etc). This can also serve as an anticounterfeiting step, and ensure medications do not get into the wrong hands. Patient adherence can be monitored by dose verification technology (eg, smart blister pack, smart pill bottle, smart autoinjector) and supplies can be reordered to ensure no interruption. Where relevant, patient training could be offered via video/voice recording, connection to a smart television, or voice commander in the home. This could also lend itself to community-based medicine through peer groups. For example, smart grids around senior living communities could facilitate group support interactions through communication with smart packaging. Such has been seen to have a positive impact on outcome measures—for example, cardiovascular disease management through group exercise routines among peers [31]. Such can promote a wellness-oriented mindset, an important behavioral driver for long-term impact. With entire populations rendered housebound due to the COVID-19 pandemic and recreational facilities closed, the notions of community wellness and maintaining physical activity have seen added interest and are likely to stay beyond the pandemic.

Deployment for Global Health

Use in developing nations, particularly rural areas where broadband networks are often unavailable, would seem an attractive proposition for use of BTM IoMT on unlicensed LPWANs. Indeed, such approaches have already been piloted in regions of sub-Saharan Africa, with successful outcomes reported [32]. Despite this validation, many studies have remained at the concept/pilot stage due to a lack of infrastructure for implementation, leading to a degree of frustration, which should be addressed as a matter of priority [32]. Pressure continues to mount, as evidenced by activities among several African nations to address mental health using mobile health tools during the COVID-19 crisis [33].

One concrete example of success was the SIMpill program trialed in South Africa. The system uses a smart drug container that sends a wireless signal when the cap is removed, signaling an adherence-related action by the patient [34]. Increased patient adherence rates ranging from 22% up to 90% were observed, suggesting fertile ground for widespread deployment [34].

In a recent study in Malawi, care for women with HIV was assessed using various technologies with a view to maximizing adherence for patients prescribed antiretroviral medications [35]. The study focused on evaluating 3 technologies for monitoring and supporting engagement in HIV care and the security screens used to validate patients' identities. Although SMS text messaging and SIM card screening were viable, biometric-based (fingerprints) ID verification showed the highest level of engagement. Drivers cited included eliminating financial barriers (costs of phone and service), security concerns (third parties gaining access to device), consistency (trading of phones/cards), and lack of device literacy [35].

Supply Chains and Clinical Trials

The studies in Malawi hint at the prospect that rural (and urban) drug dispensaries of the future might possibly function through IoMT, allowing patient verification by biometric scan through an LPWAN (or opening of smart pack/parcel pack medications delivered by drones). There seems little doubt that the technology can assist in the tracking of drug supply from manufacturer to patient and verifying adherence. The key to improved patient outcome measures lies in frequency of data capture and this should create minimal friction for the patient. As we have learned from longitudinal studies at scale, where we live affects our health and the ability to secure rich data over time is one of the keys to success in the field [36-38].

These same approaches would also benefit clinical trials being conducted in remote regions, where patient tracking and monitoring are pivotal components of trial design. Modern approaches often dictate highly segmented trials and the ability to effectively reach, engage, and monitor subpopulations is critical to addressing health disparities over time by facilitating trials with sufficient power [39].

Emergency Management

Additional opportunities abound in the form of emergency alert systems for patients. For example, lack of communications infrastructure can place pregnant mothers in a vulnerable situation in rural areas. As has been noted, makeshift communication methods used by patients in developing nations range from use of carrier pigeons to alerting passing motorists of the need for assistance, with delays leading to fatalities [40]. LPWAN could address this need and—combined with active monitoring—be used to establish a form of emergency service channel as is the case in the United States with citizens band radio channels 9 (land-based) and 16 (marine). Such might become a near-term goal for collaborations between the World Health Organization and its member nations.

There have been additional developments in this regard from the Google X team and their "Project Loon," which is developing balloons that are deployed in the stratosphere to provide wireless connectivity in remote areas around the globe. Though a current limiting factor is the 100-day service life of the balloons, they have already been deployed to assist populations in Peru impacted by floods and over Puerto Rico following Hurricane Maria, and have recently been deployed in parts of Africa [41].

Conclusion and Next Steps

The era of connected health is upon us, and the range of possibilities for deployment of IoMT devices is growing apace. Near-term expected developments that will impact global health are on the horizon. The ability to track medication shipment, delivery, and activation through smart packaging merely awaits coordination and scale.

For any systems to embed into standard practice in managed health care, they will need to be mindful of human behaviors and align with motivational factors. Groups addressing human factors engineering and studying real-world evidence will play leading roles in this, guided by insights from behavioral psychology including the models outlined by Fogg [42], Maslow [43], and Russell [44].

One can foresee devices that capture and display information changing from merely having a reporting role to having a predictive one with the benefits of artificial intelligence and machine learning trained to the user/patient. Whether data are displayed on a handheld device, through smart glasses, or on surfaces holographically [45], the utility of the data relies on contextual and relational awareness requiring careful design.

One of the early hopes for connected health in the pharmaceutical industry was the concept of the "smart pill," which could be tracked as it progressed through the gastrointestinal tract to confirm patient medication adherence. Though this remains elusive as a marketed product [46], future developments could be anticipated—for example, highly miniaturized transmitters activated when the drug encounters the low pH of the stomach or the higher pH of the ileum, sending a wireless signature directly to the cloud (or to mist or LPWAN), without the need for a smartphone [47].

Remote Health Monitoring

In addition to confirming medication dosing wirelessly, our effluent streams could become sources for rich connected health information. There is an increasing awareness of the importance of the microbiome and its implications for health and P4 medicine [48]. To capitalize on this knowledge, a properly designed "smart toilet" could analyze and track components tied to the individual (eg, as a wireless detector for the aforementioned smart pill components postactivation, or as electrochemical sensors that detect microbiota-related degradation products). There are also a number of urinalysis diagnostic kits available and FDA approved that provide colorimetric analysis for biomarker levels in effluent streams [49]. Examples include metabolites tied to diseased states, with indicator cards imaged using a smartphone and uploaded to the cloud and the diagnostics provider [50]. Such analytical systems could presumably be embedded to allow direct to cloud services from toilets, and could be demonstrated in a use case at assisted living centers. Additional rationale for this approach stems from very recent screening of effluents by regional health authorities to track SARS-CoV-2 viral spread [51].

Another possibility of direct to cloud services could be systems modeled after unmanned pop-up health clinics. Originally trialed in the United States [52,53], they have become widely adopted in China, where they were introduced to address the scale of services needed in heavily populated urban areas [54]. Biometric verification at the patient level could allow access to services through LPWAN—for example, dispensing medications, or depositing a biological sample that does not require blood draw (sweat, tears, saliva). The facility could then analyze data locally and upload to the cloud, alerting health care providers if the need for intervention arises, or confirming medication adherence for payers.

Such systems would also prove useful in the identification, tracking, and control of pandemics [51,55]. Both HIV and COVID-19 have had a severe impact in developing nations and such systems may offer a cost-effective method to provide needed services in remote areas [56]. It should also be noted that BTM applications will likely develop alongside current digital health technologies and the many groundbreaking benefits they have provided for chronic care and emergency management [57]. The possibilities are wide ranging (Figure 3) and it will be incumbent on the global health community to create solutions that will have lasting impact over time. *Think locallybutact globally* may become a mantra for this challenge.

Figure 3. Potential utilization of Beyond the Mobile technologies in global health care. EMR: electronic medical record.

Patient Treatment	Patient Security	Patient Emergencies
• Behavioral support • Adherence tracking • Remote diagnostics	• Patient ID & EMR • Smart deliveries • Anticounterfeiting	• First responder events • Vital signs alerts • Pandemic monitoring

Acknowledgments

The content is solely the responsibility of the authors and does not necessarily represent the official views of Novartis AG or Tufts.

Authors' Contributions

GBJ ideated this viewpoint, contributed to writing the manuscript, and created figures with AB. JW and AB provided subject matter expertise and contributed to writing the manuscript.

Conflicts of Interest

All authors are employees of Novartis AG and its subsidiaries. The authors declare no other competing interests.

References

1. Falch M, Henten A. Universal Service in a Digital World: The Demise of Postal Services. Nordic and Baltic Journal of Information and Communications Technologies 2018;2018(1):207-222 [FREE Full text] [doi: 10.13052/nbjict1902-097x.2018.011]
2. Lu M. Digital Divide in Developing Countries. Journal of Global Information Technology Management 2014 Sep 09;4(3):1-4 [FREE Full text] [doi: 10.1080/1097198x.2001.10856304]
3. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and Evaluating Digital Interventions to Promote Behavior Change in Health and Health Care: Recommendations Resulting From an International Workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
4. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017;3:38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]

5. The Growing Value of Digital Health. Evidence and Impact on Human Health and the Healthcare System. Iqvia Institute. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health [accessed 2022-02-02]
6. 39+ Smartphone Statistics You Should Know in 2022. Review 42. URL: https://review42.com/resources/smartphone-statistics/ [accessed 2022-02-02]
7. Zaninotto P, Batty G, Stenholm S, Kawachi I, Hyde M, Goldberg M, et al. Socioeconomic Inequalities in Disability-free Life Expectancy in Older People from England and the United States: A Cross-national Population-Based Study. J Gerontol A Biol Sci Med Sci 2020 Apr 17;75(5):906-913 [FREE Full text] [doi: 10.1093/gerona/glz266] [Medline: 31940032]
8. Vaghefi I, Tulu B. The Continued Use of Mobile Health Apps: Insights From a Longitudinal Study. JMIR mHealth uHealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
9. Smartphone users are waiting longer before upgrading — here's why. CNBC. URL: https://www.cnbc.com/2019/05/17/smartphone-users-are-waiting-longer-before-upgrading-heres-why.html [accessed 2022-02-02]
10. Chen R, Jankovic F, Marinsek N, Foschini L, Kourtis L, Signorini A. Developing Measures of Cognitive Impairment in the Real World from Consumer-Grade Multimodal Sensor Streams. 2019 Presented at: KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; August 4, 2019; Anchorage, AK. [doi: 10.1145/3292500.3330690]
11. Direct-to-Cloud Activity Working Group to Simplify Integration of Personal Health Data into Clinical Systems. Personal Connected Health Alliance. URL: https://www.pchalliance.org/news/direct-cloud-activity-working-group-simplify-integration-personal-health-data-clinical-systems [accessed 2022-02-02]
12. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019 Feb 21;2:9 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
13. Wroge T, Özkanca Y, Demiroglu C, Si D, Atkins D, Ghomi R. Parkinson's Disease Diagnosis Using Machine Learning Voice. 2018 Presented at: IEEE Signal Processing in Medicine and Biology Symposium (SPMB); 2018; Philadelphia, PA p. 1-7. [doi: 10.1109/spmb.2018.8615607]
14. Understanding the unconnected. STL Partners. URL: https://stlpartners.com/research/understanding-why-half-the-people-in-the-world-dont-have-access-to-the-internet/ [accessed 2022-02-02]
15. 2019 The State of Broadband. Broadband Commission. URL: https://www.broadbandcommission.org/publications/Pages/SOB-2019.aspx [accessed 2022-02-02]
16. Petajajarvi J, Mikhaylov K, Hamalainen M, Linatti J. Evaluation of LoRa LPWAN technology for remote health and wellbeing monitoring. 2016 Presented at: 10th International Symposium on Medical Information and Communication Technology (ISMICT); March 20, 2016; Worcester, MA. [doi: 10.1109/ismict.2016.7498898]
17. Barro P, Zennaro M, Degila J. A LoRaWAN Coverage Testbed and a Multi-optional Communication Architecture for Smart City Feasibility in Developing Countries. 2020 Presented at: International Conference on e-Infrastructure and e-Services for Developing Countries; December 3-4, 2019; Porto-Novo, Benin p. 73-83. [doi: 10.1007/978-3-030-41593-8_6]
18. Tata Communications in move to develop LoRa networks for IoT. Capacity Media. URL: https://www.capacitymedia.com/articles/3823711/tata-communications-in-move-to-develop-lora-networks-for-iot [accessed 2022-02-02]
19. Amazon Sidewalk paves the way for more connected communities. Amazon. URL: https://developer.amazon.com/en-US/blogs/alexa/device-makers/2020/09/amazon-sidewalk-paves-the-way-for-more-connected-communities [accessed 2022-02-02]
20. Semtech and YoSmart Deliver Advanced Smart Home Automation Solutions. Semtech. URL: https://www.semtech.com/company/press/semtech-and-yosmart-deliver-advanced-smart-home-automation-solutions [accessed 2022-02-02]
21. Global Smart Home Market to Hit $155 Billion by 2023. Softei. URL: https://softei.com/global-smart-home-market-to-hit-155-billion-by-2023/ [accessed 2022-02-02]
22. Health expenditure as a percentage of gross domestic product (GDP) in selected countries in 2019. Statista. URL: https://www.statista.com/statistics/268826/health-expenditure-as-gdp-percentage-in-oecd-countries/ [accessed 2022-02-02]
23. Al-Turjman F, Nawaz MH, Ulusar UD. Intelligence in the Internet of Medical Things era: A systematic review of current and future trends. Computer Communications 2020 Jan;150:644-660. [doi: 10.1016/j.comcom.2019.12.030]
24. Internet of Things in Healthcare Market. Allied Market Research. URL: https://www.alliedmarketresearch.com/iot-healthcare-market [accessed 2022-02-02]
25. MAGICSTRAP. Murata. URL: https://www.murata.com/-/media/webrenewal/campaign/ads/europe/healthcare/magicstrap_flyer_0901.ashx?la=en-gb&cvid=20150130132231181300 [accessed 2022-02-02]
26. Murata starts shipping smallest Bluetooth module in the world; reaches 500 million Bluetooth modules produced. Murata. URL: https://www.electronicspecifier.com/industries/wm/murata-lbma43pqn-bluetooth [accessed 2022-02-04]
27. The eUICC Opportunity: Harness the Power of IoT eSIMs. Sierra Wireless. URL: https://www.sierrawireless.com/resources/infographics/euicc/ [accessed 2022-02-02]
28. Yousefpour A, Fung C, Nguyen T, Kadiyala K, Jalali F, Niakanlahiji A, et al. All one needs to know about fog computing and related edge computing paradigms: A complete survey. Journal of Systems Architecture 2019 Sep;98:289-330 [FREE Full text] [doi: 10.1016/j.sysarc.2019.02.009]
29. Flores M, Glusman G, Brogaard K, Price ND, Hood L. P4 medicine: how systems medicine will transform the healthcare sector and society. Personalized Medicine 2013 Aug;10(6):565-576 [FREE Full text] [doi: 10.2217/PME.13.57] [Medline: 25342952]

30. Bittner B, Schmit Chiesi C, Kharawala S, Kaur G, Schmidt J. Connected drug delivery devices to complement drug treatments: potential to facilitate disease management in home setting. Med Devices (Auckl) 2019;12:101-127 [FREE Full text] [doi: 10.2147/MDER.S198943] [Medline: 30881151]
31. Makris E, Hu L, Jones GB, Wright JM. Moving the Dial on Heart Failure Patient Adherence Rates. Patient Prefer Adherence 2020;14:2407-2418 [FREE Full text] [doi: 10.2147/PPA.S283277] [Medline: 33324042]
32. Stephani V. Effective and needed, but not used: Why do mobile phone-based health interventions in Africa not move beyond the project status?. Berlin, Germany: TU Berlin; 2019. URL: https://depositonce.tu-berlin.de/bitstream/11303/8815/3/stephani_victor.pdf [accessed 2022-02-08]
33. Adepoju P. Africa turns to telemedicine to close mental health gap. The Lancet Digital Health 2020 Nov;2(11):e571-e572 [FREE Full text] [doi: 10.1016/S2589-7500(20)30252-1] [Medline: 33103096]
34. Broomhead S, Mars M. Retrospective return on investment analysis of an electronic treatment adherence device piloted in the Northern Cape Province. Telemed J E Health 2012;18(1):24-31. [doi: 10.1089/tmj.2011.0143] [Medline: 22150713]
35. Bengtson AM, Kumwenda W, Lurie M, Kutengule A, Go V, Miller WC, et al. Beyond mobile phones: exploring using technology to support sustained engagement in care for HIV-infected women on antiretroviral therapy. AIDS Care 2020 Aug;32(8):959-964 [FREE Full text] [doi: 10.1080/09540121.2020.1737639] [Medline: 32138524]
36. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. The Lancet 2014 Mar 15;383(9921):999-1008 [FREE Full text] [doi: 10.1016/S0140-6736(13)61752-3] [Medline: 24084292]
37. Keyes K, Galea S. Population Health Science. New York: Oxford University Press; 2016.
38. Where you live affects your health. Harvard Chan School of Public Health. 2019. URL: https://www.hsph.harvard.edu/news/hsph-in-the-news/where-you-live-affects-your-health/ [accessed 2022-02-02]
39. Clark LT, Watkins L, Piña IL, Elmer M, Akinboboye O, Gorham M, et al. Increasing Diversity in Clinical Trials: Overcoming Critical Barriers. Curr Probl Cardiol 2019 May;44(5):148-172 [FREE Full text] [doi: 10.1016/j.cpcardiol.2018.11.002] [Medline: 30545650]
40. Editorial. Reaching help on time in an emergency delivery. Safe Mother 1991(5):5-7. [Medline: 12284068]
41. Loon. Google X. URL: https://x.company/projects/loon/ [accessed 2022-02-02]
42. Toledo F, Devincenzi S, Kwecko V, Mota F, Botelho S. A framework for modeling Persuasive Technologies based on the Fogg Behavior Model. 2018 Presented at: IEEE Frontiers in Education Conference (FIE); October 3, 2018; San Jose, CA p. 1-5. [doi: 10.1109/fie.2018.8659195]
43. Kenrick D, Griskevicius V, Neuberg S, Schaller M. Renovating the Pyramid of Needs: Contemporary Extensions Built Upon Ancient Foundations. Perspect Psychol Sci 2010 May;5(3):292-314 [FREE Full text] [doi: 10.1177/1745691610369469] [Medline: 21874133]
44. Open Culture. URL: https://www.openculture.com/2017/08/bertrand-russell-reveals-the-4-human-desires-that-make-our-world-acquisitiveness-rivalry-vanity-love-of-power.html [accessed 2022-02-02]
45. Hands-on: Looking Glass Portrait is a window into the holographic data era. Venturebeat. URL: https://venturebeat.com/2021/01/04/hands-on-looking-glass-portrait-is-a-window-into-the-holographic-data-era/ [accessed 2022-02-02]
46. ABILIFY MYCITE. Abilify. URL: https://www.abilifymycite.com/ [accessed 2022-02-02]
47. Kong YL, Zou X, McCandler CA, Kirtane AR, Ning S, Zhou J, et al. 3D-Printed Gastric Resident Electronics. Adv Mater Technol 2019;4(3):1800490 [FREE Full text] [doi: 10.1002/admt.201800490] [Medline: 32010758]
48. Peñalver Bernabé B, Cralle L, Gilbert JA. Systems biology of the human microbiome. Curr Opin Biotechnol 2018 Jun;51:146-153. [doi: 10.1016/j.copbio.2018.01.018] [Medline: 29453029]
49. Healthy.io. URL: https://healthy.io/ [accessed 2022-02-02]
50. At-Home UTI Test Kit. Testcard. URL: https://testcard.com/ [accessed 2022-02-02]
51. Larsen DA, Wigginton KR. Tracking COVID-19 with wastewater. Nat Biotechnol 2020 Oct;38(10):1151-1153 [FREE Full text] [doi: 10.1038/s41587-020-0690-1] [Medline: 32958959]
52. HealthSpot assets snapped up by former customer Rite Aid for $1.15M. Mobihealth News. URL: https://www.mobihealthnews.com/content/healthspot-assets-snapped-former-customer-rite-aid-115m [accessed 2022-02-02]
53. Higi. URL: https://secure.higi.com/locator [accessed 2022-02-02]
54. China Now Has Non-Staffed Minute Clinics and Pharmacies – Time to Raise Your Innovation Goals? Useful Arts. URL: http://usefularts.us/2018/11/23/china-no-staff-minute-clinic/ [accessed 2022-02-02]
55. Martin T, Karopoulos G, Hernández-Ramos J, Kambourakis G, Nai Fovino I. Demystifying COVID-19 Digital Contact Tracing: A Survey on Frameworks and Mobile Apps. Wireless Communications and Mobile Computing 2020 Oct 17;2020:1-29 [FREE Full text] [doi: 10.1155/2020/8851429]
56. Ataguba OA, Ataguba JE. Social determinants of health: the role of effective communication in the COVID-19 pandemic in developing countries. Glob Health Action 2020 Dec 31;13(1):1788263 [FREE Full text] [doi: 10.1080/16549716.2020.1788263] [Medline: 32657669]

57. Winders WT, Garbern SC, Bills CB, Relan P, Schultz ML, Trehan I, et al. The effects of mobile health on emergency care in low- and middle-income countries: A systematic review and narrative synthesis. J Glob Health 2021 Apr 03;11:04023 [FREE Full text] [doi: 10.7189/jogh.11.04023] [Medline: 33828846]

Abbreviations

BTM: Beyond the Mobile
IoMT: Internet of Medical Things
LoRaWAN: long-range wide area network
LPWAN: low-power wide area network

Edited by G Eysenbach; submitted 08.06.21; peer-reviewed by B Mesko, V Gupta; comments to author 21.07.21; revised version received 22.07.21; accepted 20.12.21; published 11.02.22.

Please cite as:
Jones GB, Bryant A, Wright J
Democratizing Global Health Care Through Scalable Emergent (Beyond the Mobile) Wireless Technologies
JMIR Biomed Eng 2022;7(1):e31079
URL: https://biomedeng.jmir.org/2022/1/e31079
doi:10.2196/31079
PMID:38875689

Original Paper

The Cole Relaxation Frequency as a Parameter to Identify Cancer in Lung Tissue: Preliminary Animal and Ex Vivo Patient Studies

Les Bogdanowicz[1,2], PhD; Onur Fidaner[1], PhD; Donato Ceres[1], PhD; Alexander Grycuk[1], BS; Martina Guidetti[1], PhD; David Demos[3,4], MD

[1]Novascan Inc, Chicago, IL, United States

[2]Department of Biomedical Engineering, University of Illinois at Chicago, Chicago, IL, United States

[3]Aurora St. Luke's Medical Center, Milwaukee, WI, United States

[4]Department of Cardiothoracic Surgery, Aurora Healthcare, Milwaukee, WI, United States

Corresponding Author:
Les Bogdanowicz, PhD
Novascan Inc
Chicago, IL
United States
Phone: 1 3129145370
Email: les@novascanllc.com

Abstract

Background: Lung cancer is the world's leading cause of cancer deaths, and diagnosis remains challenging. Lung cancer starts as small nodules; early and accurate diagnosis allows timely surgical resection of malignant nodules while avoiding unnecessary surgery in patients with benign nodules.

Objective: The Cole relaxation frequency (CRF) is a derived electrical bioimpedance signature, which may be utilized to distinguish cancerous tissues from normal tissues.

Methods: Human testing ex vivo was conducted with NoduleScan in freshly resected lung tissue from 30 volunteer patients undergoing resection for nonsmall cell lung cancer. The CRF of the tumor and the distant normal lung tissue relative to the tumor were compared to histopathology specimens to establish a potential algorithm for point-of-care diagnosis. For animal testing in vivo, 20 mice were implanted with xenograft human lung cancer tumor cells injected subcutaneously into the right flank of each mouse. Spectral impedance measurements were taken on the tumors on live animals transcutaneously and on the tumors after euthanasia. These CRF measurements were compared to healthy mouse lung tissue. For porcine lung testing ex vivo, porcine lungs were received with the trachea. After removal of the vocal box, a ventilator was attached to pressurize the lung and simulate breathing. At different locations of the lobes, the lung's surface was cut to produce a pocket that could accommodate tumors obtained from in vivo animal testing. The tumors were placed in the subsurface of the lung, and the electrode was placed on top of the lung surface directly over the tumor but with lung tissue between the tumor and the electrode. Spectral impedance measurements were taken when the lungs were in the deflated state, inflated state, and also during the inflation-deflation process to simulate breathing.

Results: Among 60 specimens evaluated in 30 patients, NoduleScan allowed ready discrimination in patients with clear separation of CRF in tumor and distant normal tissue with a high degree of sensitivity (97%) and specificity (87%). In the 25 xenograft small animal model specimens measured, the CRF aligns with the separation observed in the human in vivo measurements. The CRF was successfully measured of tumors implanted into ex vivo porcine lungs, and CRF measurements aligned with previous tests for pressurized and unpressurized lungs.

Conclusions: As previously shown in breast tissue, CRF in the range of 1kHz-10MHz was able to distinguish nonsmall cell lung cancer versus normal tissue. Further, as evidenced by in vivo small animal studies, perfused tumors have the same CRF signature as shown in breast tissue and human ex vivo testing. Inflation and deflation of the lung have no effect on the CRF signature. With additional development, CRF derived from spectral impedance measurements may permit point-of-care diagnosis guiding surgical resection.

(JMIR Biomed Eng 2022;7(1):e35346) doi:10.2196/35346

KEYWORDS

Cancer; Detection; Lung Cancer; Non Small Cell Lung Cancer; Spectral Impedance; Cole Relaxation Frequency; biomedical engineering; medical device; scan; testing; identify; lungs; respiratory; diagnosis; model; animal testing; in-vivo; human testing

Introduction

Background

Lung cancer is the cause of 20% of all cancer-related deaths globally and is the most commonly diagnosed cancer worldwide, representing 13% of all cancer diagnoses. Globally it is estimated that lung cancer accounts for 2 million cases per year and 1.7 million deaths per year [1]. The economic health burden of lung cancer is estimated to be 1.5% of the world's gross domestic product, approximately 1.2 trillion dollars [2].

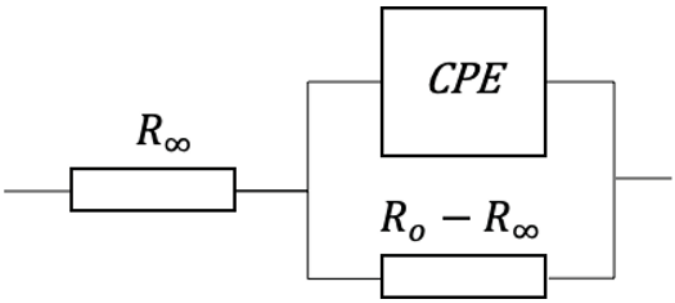
The diagnosis of lung cancer begins with the detection of pulmonary nodules or masses using chest X-rays and low-dose computed tomography (CT) technologies. Nodules and/or masses appear in 1 out of every 500 chest X-rays, and only about 40% are cancerous [3]. These nodules usually develop in the deep periphery of the lung, requiring navigation to branching airway structures to reach them. The nonsurgical gold standard technique to differentiate benign from malignant pulmonary nodules is either a transthoracic needle biopsy or bronchoscopy biopsies [4]. These techniques report a diagnostic accuracy below 78% and 60%, respectively, with decreasing sensitivity for smaller nodules [4,5]. Consequently, when small nodules (<1 cm) are detected, they are typically monitored for 3 to 6 months and are pursued upon growth. More advanced technologies that can detect cancer accurately in small malignant lung nodules can help achieve early diagnosis and prompt treatment, which is critical to lung cancer survival. In recent years low-dose CT screening was shown to reduce lung cancer mortality by shifting cancer detection from symptomatic late-stage cancers to early-stage cancers, supported by the National Lung Screening Trial and the NELSON trial [6]. The relative survival rate for cancerous nodules is about 50% after 5 years, rising to 80% if the nodule is first diagnosed at a diameter of 1 cm or less [3]. In this paper, the authors propose a novel technology, NoduleScan, that provides an accurate diagnosis of cancerous lung nodules via an electrical assessment of the tissue. The accuracy of NoduleScan will be tested by comparing the outcomes (cancerous vs noncancerous) obtained with the new technology to the results of the pathology report on 3 different tissues: (1) human lung tumors from 30 patients (human lung testing ex vivo); (2) A549 human lung tumor xenograft grown on mouse model and noncancerous lung tissue from the same mouse model, tested both in vivo, to include the effect of perfusion, and ex vivo (animal testing); and (3) healthy ventilated porcine lungs with A549 cancerous nodules insertion to include the effect of breathing and to demonstrate detection sensitivity (porcine lung testing). We hypothesized that NoduleScan would discern between cancerous and noncancerous, as was previously shown in breast and skin [7-9].

NoduleScan Technology

Bioimpedance Modeling for Cancer Detection

The characterization of the bioelectrical properties of tissues is often achieved using electrical circuits models, which attempt to describe in electrical terms the frequency-dependent behavior of ion movement within the extracellular and intracellular media and across the cellular membrane, as well as membrane polarization effects. The Cole-Cole impedance model (Figure 1) is often used to describe the impedance behavior of biological tissue [10].

Figure 1. Equivalent circuit model. CPE: constant phase element.



In this representation, the constant phase element (CPE) is an empirical circuit element purely devised to describe polarization phenomena as a result of system inhomogeneity:

where C_α is a pseudo-capacitance, and α is its order ($0<\alpha<1$).

The Cole-Cole equation (2) derives in principle from the CPE model, specifically when describing impedance behavior of biological tissue as

Where $\tau=1/2\pi f_c$ is the relaxation constant, $C_\alpha=\tau^\alpha/(R_0-R_\infty)$, and f_c is the Cole relaxation frequency (CRF) [11].

The fundamental parameters (R_0, R_∞, and f_c) can be used to describe and characterize the bioimpedance behavior of a tissue. Although these parameters do not necessarily describe the physical mechanisms involved in the impedance measurement, in the literature, R_0 and R_∞ have been associated with the

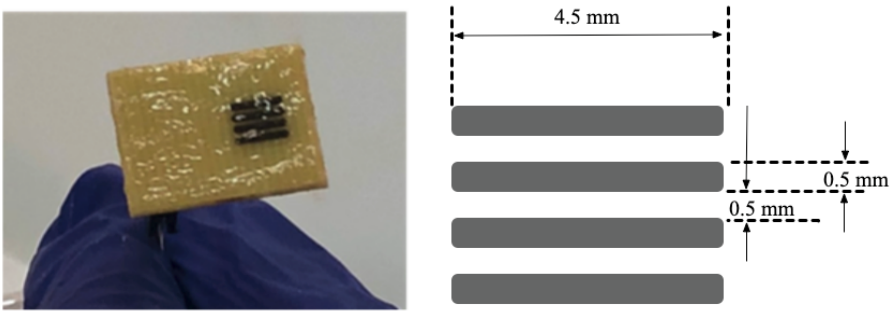
extracellular medium and fluids, while α and C_α with membrane polarization and ionic permeability [12-14]. Accordingly, we have found that only the CRF might be sufficient to discriminate between cancerous and noncancerous tissue in the frequency range from 10^3 Hz to 10^7 Hz, often referred to as the β dispersion region, where cellular membranes have the strongest frequency response [8]. This enhanced dispersion sensitivity of the cellular membrane in the β region might provide the physiological interpretation of why the differential membrane polarization and ionic permeability between healthy and cancerous cells in a given tissue result in distinct impedance signatures, and more particularly of CRF values [15].

Recently, it has been demonstrated in a cohort of 373 patients that a single parameter, CRF, can be used to detect breast cancer with sensitivity and specificity of 95% and 87%, respectively [8]. Similarly, CRF has been used to identify skin cancer with sensitivity and specificity of 100% and 90%, respectively [7,9]. In the same way, the CFR parameter will be used in the present study as a biomarker for lung cancer detection, based on impedance spectroscopy in the β frequency range [15].

NoduleScan Device Used for Testing

The CRF measurements are obtained using a tetrapolar electrode system, as shown in Figure 2. The 4 electrodes are prepared by electroplating platinum black (PtBlk), which results in nonpolarizable electrode surfaces suitable for bioimpedance measurement of tissue. An electrolytic gel [16] is used to provide an electrical coupling between the electrodes and the tissue. When in contact with the tissue, a spectral impedance measurement is recorded at discrete frequencies (32 frequencies per decade) in the β frequency region. In this tetrapolar configuration, the two outer electrodes are the stimulating electrodes that drive a current through the tissue by applying a $1V_{rms}$ voltage, while the two inner electrodes are used to measure the voltage differential. Amplitude and phase of complex bioelectrical impedance are then calculated from the ratio of the voltage differential to the current. The impedance measurement is repeated for each discrete frequency in the β frequency region. Our testing apparatus is optimized to have the highest fidelity in the $2x10^3$Hz-$1x10^7$ Hz frequency range.

Figure 2. Photograph and dimensions of measuring electrodes.



Analysis of Impedance Data

The raw impedance data is analyzed through the following steps.

1. Remove parasitic resistance-capacitance artifacts coming from the reactive circuit elements in the probe and the cable.
2. Nyquist path analysis of data for initial determination of cancer
3. Find α and f_c values that result in the highest correlation between the impedance formula in equation (1) and the data.
   - If the CRF value (f_c) is in the range of 10^5 Hz-2.1 x 10^6 Hz (defined as the cancerous range) [8], then the scan is characterized as "cancer," otherwise "no cancer."
   - If there is no CRF value (f_c) present, then the scan is characterized as "no cancer," solely based on Nyquist path analysis results.
4. Typically, multiple scans are taken from each sample at slightly different locations. If at least one scan is cancer, then the sample is characterized as "cancer." Otherwise, if all scans from a sample are "no cancer," then the sample is characterized as "no cancer."

As an example of the application of the NoduleScan technology, Figure 3 shows the imaginary component of complex impedance computed for two specimens obtained from a mouse model. The plots show the imaginary component of measured impedance following removal of parasitic capacitance components. Figure 3A shows the impedance response of an A549 tumor grown on a mouse, which was confirmed to be "cancer" via pathology. For this sample, our scans detected clear peaks at about $6x10^5$ Hz, categorizing it as "cancer." Figure 3B shows the impedance response of a healthy lung sample extracted from the same mouse, which was confirmed to be "no cancer" via pathology. For this sample, no peaks could be detected in the cancer range; hence the sample was categorized as "no cancer." Figure 4 shows the Nyquist path plots of the same data set in Figure 3 for data points in the cancer frequency range of $1x10^5$-$2.1x10^6$ Hz. Cancer scans tend to have a concave arch shape, as suggested by the Cole function (Eq.1 and is the basis for initial cancer/no cancer classification in Step 2.

Figure 3. Sample of spectral impedance measurements obtained from mouse model tests. Multiple curves represent repeated measurements of the same tissue in slightly different locations. (a) A549 tumor extracted from mouse (b) normal lung tissue extracted from mouse.

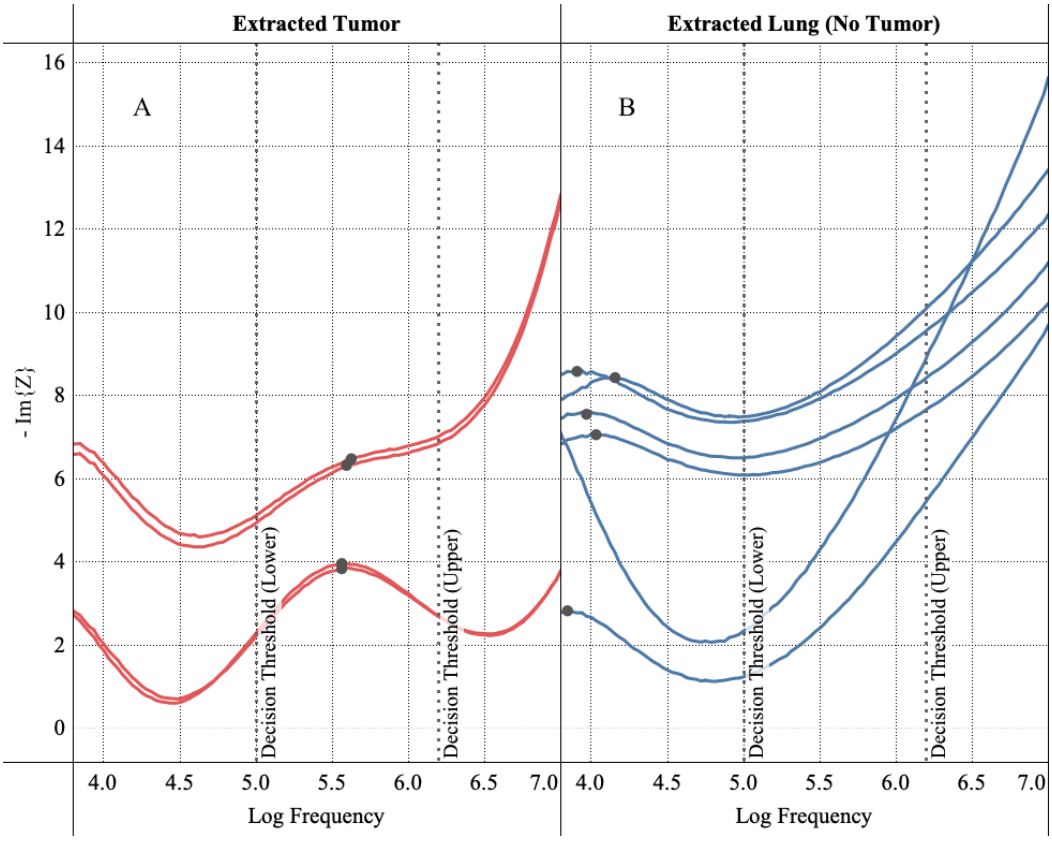
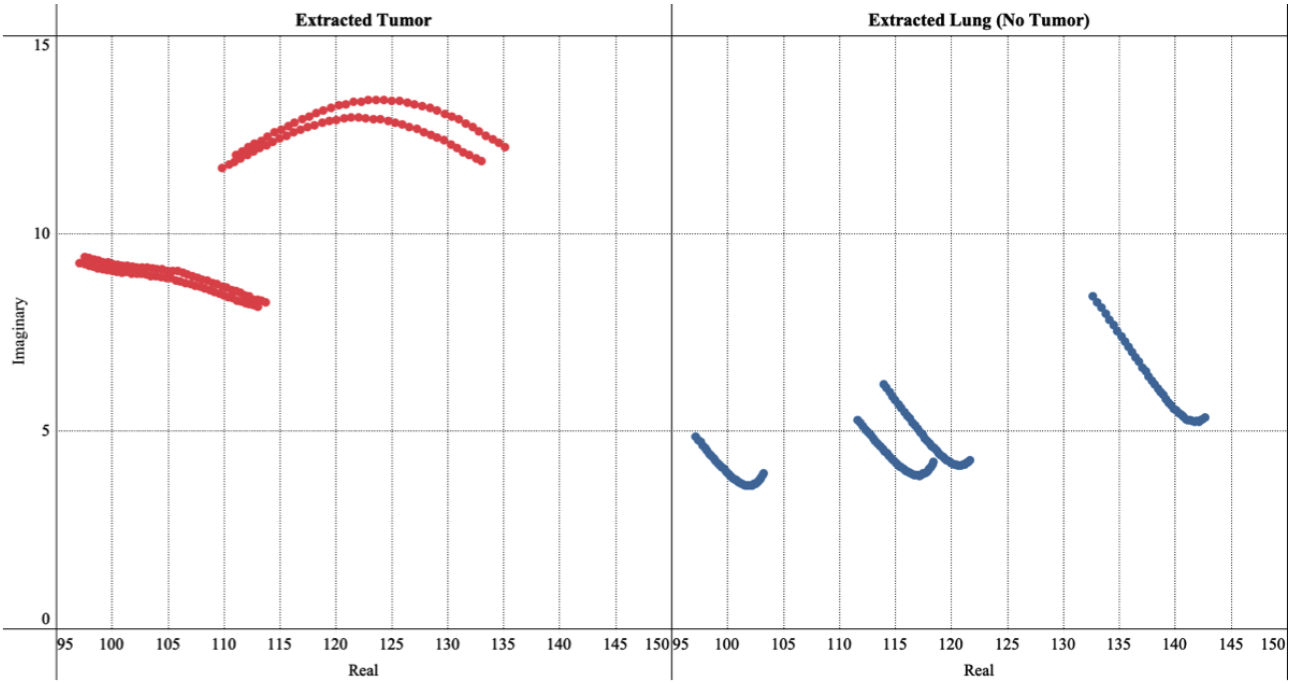


Figure 4. Nyquist path plot, which shows the imaginary part of measured impedance vs the real part for data obtained from mouse model tests, for frequencies in the cancer range $1x10^5$ Hz - $2.1x10^5$ Hz.



Methods

Human Lung Testing Ex Vivo

Patients (n=30) were recruited by informed consent from the partner Aurora St. Luke's Medical Center in Milwaukee, Wisconsin, under the approval from WCG IRB (study number: 1264221). Research subjects were recruited from patients at any site within the Aurora Healthcare system, who had a diagnosis of nonsmall cell lung carcinoma (NSCLC) with tumor size >1.4 cm, and were elected for a surgical wedge

segmentectomy or lobectomy procedure with or without lymph node dissection.

- *Study population:* Adult men and women presenting with a diagnosis of NSCLC who were candidates for wedge resection or anatomic resection, or thoracic nodal dissection.
- *Inclusion criteria:* Adult male and female patients, 18 years of age or greater, with a biopsy-proven diagnosis of lung cancer and who are scheduled for a wedge resection or anatomic resection. Preoperative diagnosis is NSCLC that may include the following: invasive carcinoma, squamous cell carcinoma, or adenocarcinoma for those patients scheduled for surgical excision of tissue.
- *Exclusion criteria:* Patients who refuse consent for the study of their excised tissue, patients with a diagnosis of small cell lung cancer, patients with a diagnosis of NSCLC who underwent induction therapy with either chemotherapy or radiation, and patients who are not candidates for surgical resection.

Research subjects were preoperatively consented to by the surgeon and Aurora research coordinator. The patient then underwent a lobectomy or segmentectomy as per standard of care, and the excisions were bisected. The authors collaborated with the pathology staff to select tissue from two sites for each patient: a piece from the suspected cancerous region and a piece from noncancerous tissue at least 5 cm away from the suspected tumor site. CRF measurements were performed by placing the electrode in different positions on each tissue sample (tumor and away-from-tumor). Once spectral impedance scanning was completed (multiple scans for both tumor and away-from-tumor pieces), the tissue samples were returned to the pathology staff for immediate processing and histological evaluation.

Animal Testing

Overview

With encouraging results from the human ex-vivo lung assessment, we pursued answering questions on the ability to discern between cancer and noncancer in perfused tissue. Twenty mice were inoculated with an NSCLC line, and all survived to full tumor growth. CRF measurements were taken on tumors through the skin with a live animal, with the tumor exposed after euthanizing the animal, and subsequently, the tumor itself was excised and measured. These results were compared to measurements on the excised healthy lung tissue from the same mouse.

Bagg Albino inbred nude mice BALB/c NU/NU (n=25) were implanted with tumor cells that were grown in culture (A549 lung sarcoma cells obtained from American Type Culture Collection, ATCC: CCL-185). These A549 cells (2 × 106 cells in approximately 100 µL sterile PBS) were injected subcutaneously into the right flank of each mouse.

Animals with tumor sizes between 0.1 cm^3 to 1.0 cm^3 were utilized in this study. Tumor measurements were monitored twice weekly until termination. Tumor size was calculated using the formula: volume=½ Length x $width^2$.

Animal testing was performed at the University of Illinois (UIC) Toxicology Research Laboratory (TRL). The protocol for the study was approved by the UIC Office of Animal Care and Institutional Biosafety (approval number: 20-147).

Spectral Impedance Measurements

1. *Impedance measurements on tumors with skin intact on live mice:* CRF was measured on live mice on the surface of the intact skin at the tumor site. The live mice were held steady by the nape of their neck and tail, and measurements were taken on the tumor site holding the electrodes in place. Upon completion of this test, the mice were euthanized.
2. *Impedance measurements on dead mice with exposed tumor:* The tumor site was exposed by cutting a flap from the posterior end of the mouse, extending anteriorly until the entirety of the tumor was accessible. The measurement electrode was then placed directly on the exposed tumor, and multiple spectral impedance scans were recorded.
3. *Impedance measurements on ex vivo tumor:* The tumor was excised in its entirety from the animal and bisected. The sliced faces of the tumor pieces were then placed on the measurement electrode one by one, and spectral impedance scans were taken from each piece. One piece was sent to histology for evaluation.
4. *Impedance measurements on noncancerous excised mice lung:* The lungs (noncancerous tissue) were excised from the animal and sliced so that the interior faces were exposed. Spectral impedance measurement scans were recorded. One piece was sent to histology for evaluation.

Porcine Lung Testing

Overview

Two healthy porcine lungs were received within 24 to 26 hours of euthanization with the trachea and vocal cords intact and shipped to the research facility overnight at a constant temperature of 4°C. The vocal cords were cut out, and a manual ventilator was attached and secured tightly in place. The lungs were inflated and deflated using the ventilator, and the applied pressure was measured using a gauge.

Separately, 5 human lung tumors (A549 cell line) were grown in 5 separate mice following the procedure in the murine xenograft cancer method discussed earlier. Following tumor inoculation in mice, tumor size was monitored regularly. Once tumor size exceeded 500 mm^3, the mouse was euthanized, and the tumor was excised and bisected. Consequently, each bisected piece size was greater than 250 mm^3 on average and ready to be inserted in the porcine lung tissue in order to undergo spectral measurements during simulated breathing.

Spectral Impedance Measurements

1. *Impedance measurements on ex vivo tumor:* Spectral impedance measurements were recorded on A549 tumors immediately after excision and bisection. The measurements were done ex vivo on the inner bisected surface of the tumors. One bisected piece was placed in 10% formalin and sent to a histology laboratory for evaluation.
2. *Impedance measurements on healthy porcine lung tissue:* Spectral impedance measurements were performed on porcine lungs when the lung was deflated. Subsequently, the lung was pressurized to 20 mmHg and kept at an inflated

state with constant pressure. Lastly, the air was slowly released to deflate the lung at a rate of <1 mmHg per second. During the deflation process, multiple spectral impedance scans were performed.

3. *Impedance measurements on porcine lung with inserted tumor:* A flap was created on the lung to expose the underlying subsurface lung tissue by cutting the outermost pleural layer of the lung. The flap tissue was cut to a width that matched the size of the A549 tumor previously excised from the mouse in order to ensure the tumor fit in the lung subsurface. The thickness of the flap was measured and recorded using a caliper; the target cut tissue thickness was <300 microns. The bisected A549 tumors were placed in the created lung subsurface with the flap covering the tumor. Measurements were taken on the surface of the lung at the tumor location, with the lungs in an inflated and deflated state. This was repeated for each of the tumor samples. Saline solution was applied to the lungs every 15 min during the studies to maintain tissue hydration.

Results

Human Lung Testing Outcomes

Patients' demographic and tumor characteristics are presented in Table 1.

The maximum tumor size presented was 7.5 cm. Per the protocol, subjects with tumor sizes larger than 1.4 cm were included in the study. The mean size of the tumors was 3.22 cm and median 3.0 cm, with a standard deviation of 1.47 cm. Figure 5 top plot shows a histogram of CRF values cancerous tumor samples, as determined by pathology. The bottom plot shows the CRF measurements determined to be healthy tissue by the pathologist. Note that multiple scans were taken per sample. When cancer scans are present, a sample is categorized as "cancer," and below threshold scans are ignored. The data points in the histogram represent the log-mean of measured CRF values.

Cancerous regions for CRF values ($1x10^5$ Hz-$2.1x10^6$ Hz) were compared to the pathology report for cancerous tissue. One tumor sample (confirmed to be "cancer" via pathology) did not yield a CRF value in the cancerous frequency range, resulting in a false negative. We found CRF values in the cancerous range for three 'normal' tissue specimens (confirmed to be "no cancer" via pathology) that were at least 5 cm away from the tumor lesion, resulting in false positives. The overall sensitivity and specificity values (combined tumor and distant tumor tissue) are 97% and 87%, respectively. The positive predictive value is 88%, and the negative predictive value is 96%. Results are summarized in Table 2.

Table 1. Patients' demographic and tumor characteristics.

Characteristics	No. of patients
Patient age group (54-85 years)	30
Males	16
Females	14
Adenocarcinoma	21
Squamous Cell Carcinoma	9
pN0	27
pN2	3
ypT0	1
pT1a	0
pT1b	7
pT1c	7
mpT1c	1
pT2	1
pT2a	4
mpT2a	1
pT2b	3
pT3	3
pT4	2

Figure 5. Histogram of Cole relaxation frequency (CRF) measurements from human lung samples.

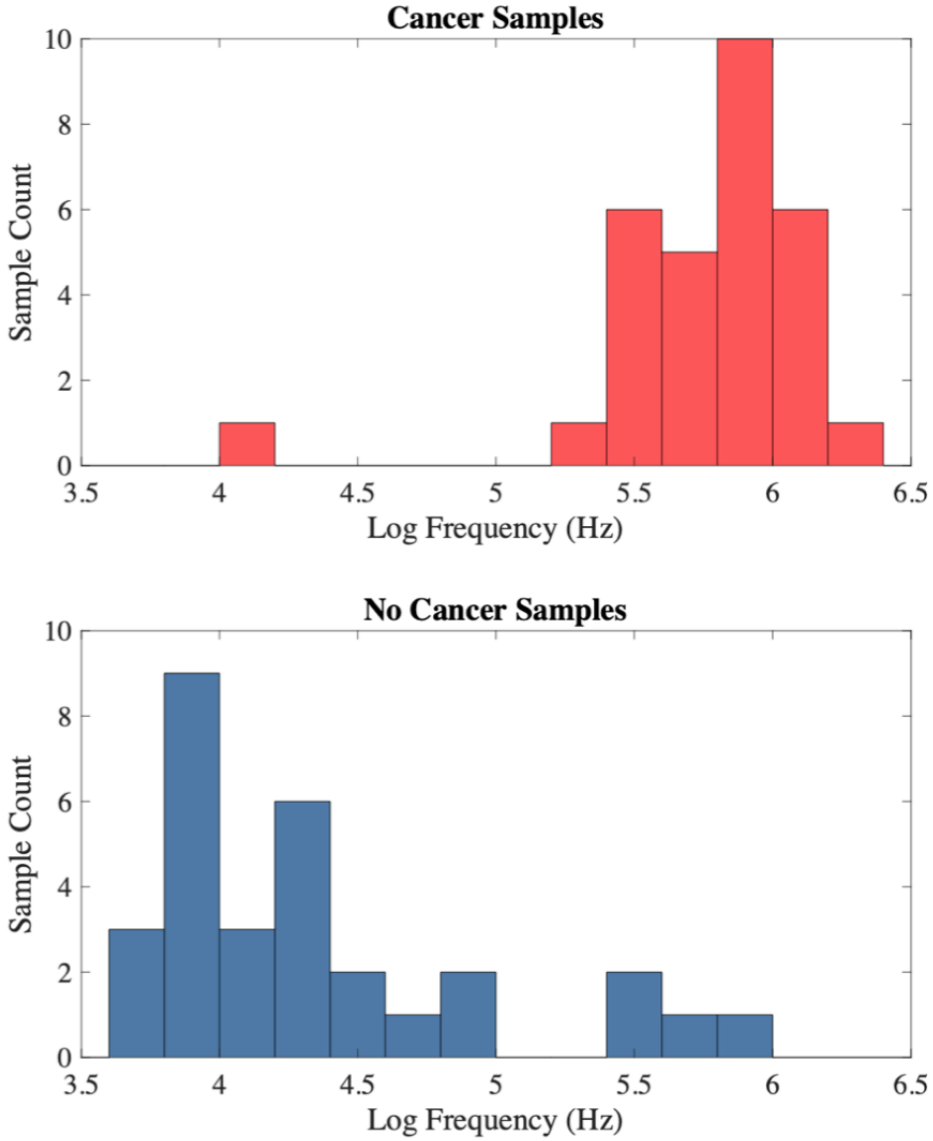


Table 2. Sensitivity and specificity of human lung testing.

	Pathology outcome tumor tissue		Pathology outcome distant tissue (normal)		Overall	
	Cancer	No cancer	Cancer	No cancer	Cancer	No cancer
CRF[a] DX[b]						
Cancer	28	0	1	4	29	4
No Cancer	1	1	0	25	1	26

[a]CRF: Cole relaxation frequency.

[b]DX: diagnosis.

Mouse Model Testing

- *Tumor growth:* Tumor height and diameter were measured for each mouse in regular time intervals following inoculation. The eventual tumor sizes had a mean of 432 mm^3 and a standard deviation of 222 mm^3.
- *Impedance measurements on tumors with skin intact on live mice:* CRF peaks were detected in the cancer range for all samples confirmed as tumors from the pathology report. The median CRF value was determined to be $3.11x10^5$ Hz. In some samples, multiple distinct CRF peaks can be observed—a phenomenon reported earlier in the literature

[8]. In tumor samples, we typically observed a primary CRF in the higher frequency range, the cancerous range, and a secondary CRF value in the lower frequency range. These secondary CRF values correspond to noncancerous tissue. The presence of noncancerous CRF measurements and cancerous CRF measurements reflect the heterogeneity of lung cancer [17].

- *Impedance measurements on dead mice with exposed tumor:* Impedance measurements were recorded on various spots on the exposed tumor. The median CRF value was determined to be 4.36×10^5 Hz.
- *Impedance measurements on extracted tumor:* Measurements of the sliced tumor were similar to those of the exposed tumor. The median CRF value of sliced tumors was determined to be 8.32×10^5 Hz. All histology data confirmed that the tissue pieces were cancerous. This offers a direct comparison of the CRF curves for tumor and no-tumor cases. A comparison of CRF frequency histograms for tumor slices and lung slices is provided in Figure 6 below. The horizontal axis represents the base 10 logarithms of frequency (Hz) values. The histograms are formed using CRF values from live mice, exposed tumor, extracted tumor, and extracted lung (healthy tissue). Typically, 3-4 scans are taken from each sample. When cancer scans are present, a sample is categorized as "cancer," and below threshold scans are ignored. The data points in the histogram represent the log-mean of measured CRF values.

The analysis reveals that all the CRF measurements for lung slices (no-tumor sample) are below the decision threshold of 1×10^5 Hz, with a median CRF frequency of 9.68×10^3 Hz. On the other hand, most of the CRF peaks for tumor slices samples are above the decision threshold, with a median CRF frequency of 8.28×10^5 Hz.

In Table 3, we show the comparison of pathologist outcome of tumor samples compared to CRF scans and healthy lung tissue compared to CRF measurements. For the 10 samples analyzed, the outcomes corresponded to 100% sensitivity, 100% specificity, and 100% positive predictive value and negative predictive value.

Figure 6. Histogram of Cole relaxation frequency (CRF) measurements of A549 tumors grown on mice.

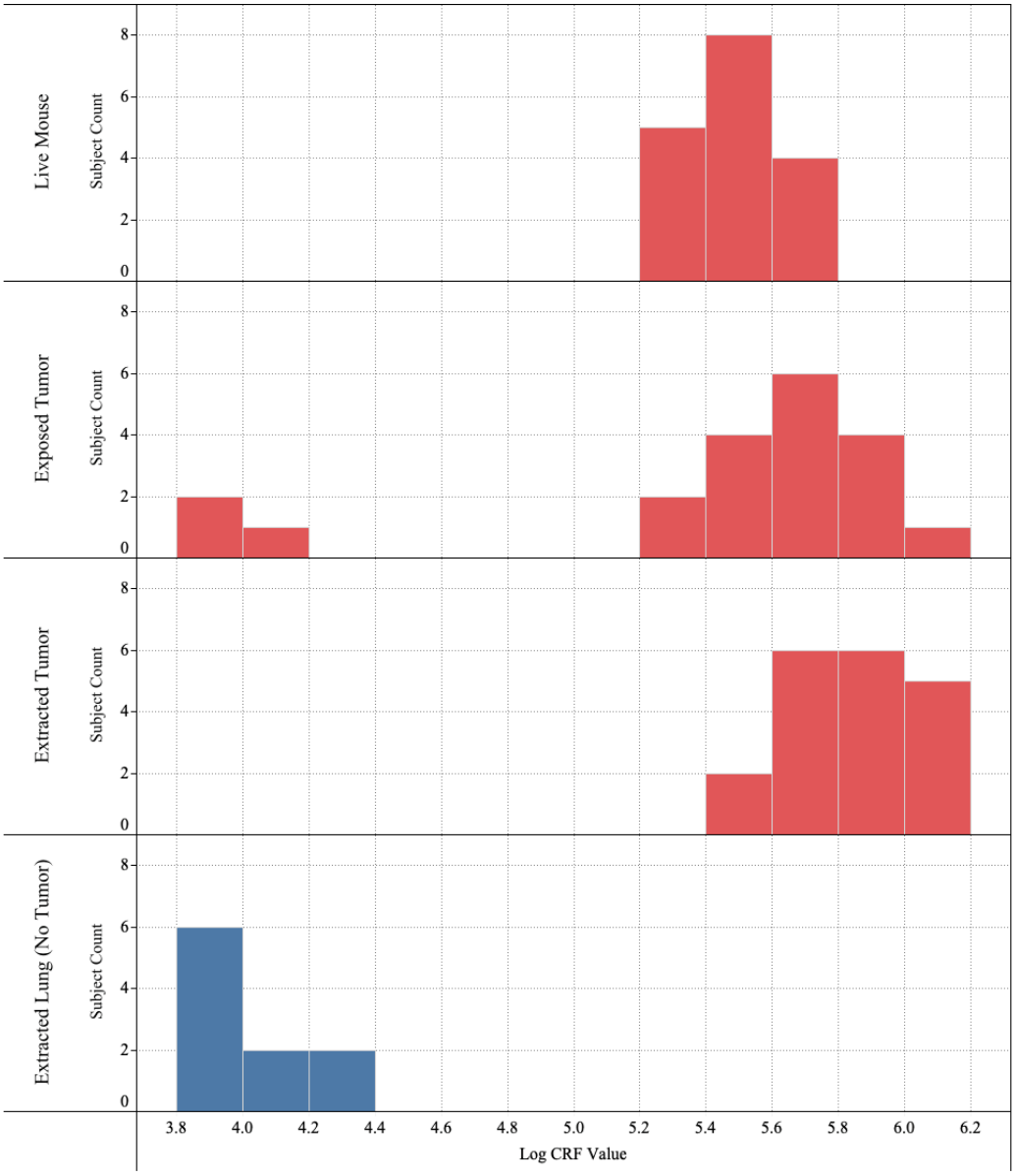


Table 3. Sensitivity and specificity of mouse model testing.

	Pathology outcome tumor tissue		Normal lung tissue (no cancer)		Pathology outcome overall	
	Cancer	No cancer	Cancer	No cancer	Cancer	No cancer
CRF[a] DX[b]						
Cancer	10	0	0	0	10	0
No cancer	0	0	0	10	0	10

[a]CRF: Cole relaxation frequency.

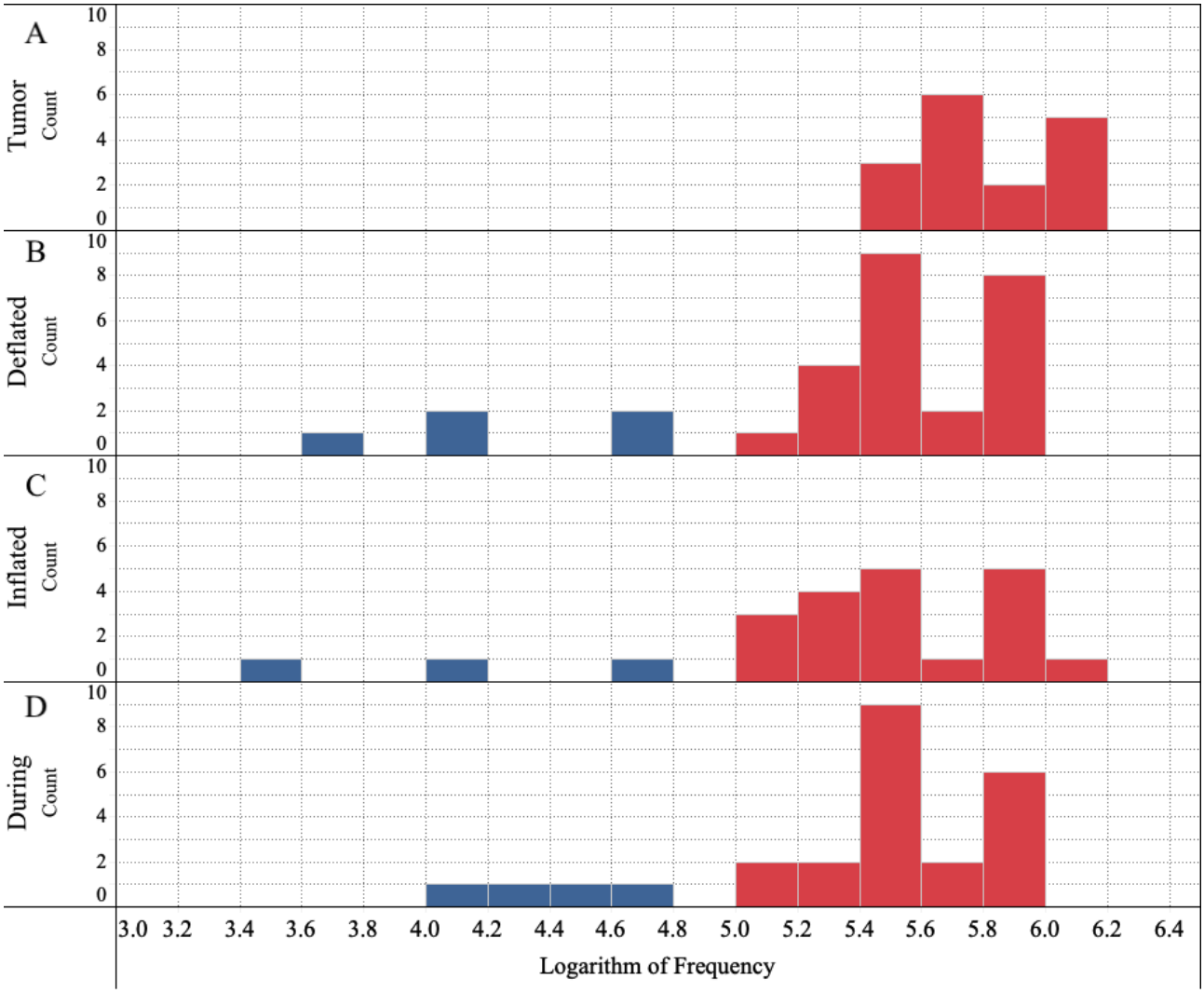
[b]DX: diagnosis.

Porcine Lung Testing

Impact of Simulated Breathing on CRF Values

To understand the impact of simulated breathing, in Figure 7, we compare the histograms for deflated, inflated, and during simulated breathing cases. Data include subsurface (electrode on flap) and subsurface (electrode on tumor) cases combined. A no CRF peak measured is classified as noncancer. There is a shift in the log-mean of CRF peak frequencies among simulated breathing scenarios (deflated/inflated/during) and is within 0.05 log cycles. There is a smaller amount of CRF measurements in the no cancer range resulting in limited information in histogram plots.

Figure 7. Histograms of Cole relaxation frequency (CRF) data for (A) Tumor Only, (B) Deflated, (C) Inflated, and (D) During Simulated Breathing cases.



Discussion

Principal Findings

We set out to test a novel technology, NoduleScan, to provide a diagnosis of cancerous lung nodules via an electrical assessment of the tissue. Further, we have shown that the diagnosis of cancerous and noncancerous tissue is in line with those seen in breast tissue and skin tissue [7-9].

We have shown that ex vivo CRF measurements of excised human lung cancerous nodules can differentiate cancerous and noncancerous tumors with a sensitivity of 97% and specificity of 87%.

In xenograft mouse models, tumor growth was found to be in agreement with published literature on mice inoculated with the same amount of A549 cells [18]. A549 lung sarcoma tumors grown in mice exhibit the same distinction in CRF values as ex vivo lung tissues and those of previously reported breast and skin tissue. The measurements remain the same when measured in vivo through the skin, with the skin opened tumor exposed and tumor excised. There was no difference in CRF measurements with the tumor perfused or not perfused. When the tumor measurements, confirmed to be cancerous by a pathological examination, were compared to the healthy lung tissue of the sacrificed mouse, the sensitivity and specificity were 100%.

Clear separation of CRF distributions, as shown in Figure 6, demonstrates that CRF is a viable parameter to identify the presence of cancer. Cancer determination is done through multiple scans from the same tissue sample. For each of these below-threshold primary peaks, there was at least one other scan from the same tissue that exhibited cancer-range CRF frequency, hence correctly classifying the sample as “cancer.”

CRF values could be detected and remained relatively unchanged during simulated breathing. CRF values are in the same ranges as shown in ex vivo human tumor tests, xenograft mice, and previously demonstrated breast and skin tissues.

Conclusions

Patients of Interest

Table 2 shows a pathological assessment of a suspect lesion as noncancerous and, secondarily, a normal tissue assessment as cancerous. These patients were patient #6036, a 60 year old female undergoing a right upper lobectomy, and study patient #6086, a 72 year old female undergoing a right lower lobectomy. For patient #6036, the nodule presented did not show any CRF values in the cancerous range, and this was later confirmed by pathological examination as fibrotic tissue, TNM status ypT0, pN0, and no cancer was detected showing concordance with the scanned results. Patient #6086 measured CRF values in the cancerous range for tissue distant from the tumor (>5 cm,) and pathological examination confirmed positive margins and thus showed concordance with our measurements.

NoduleScan Technology

The NoduleScan assessment is completed rapidly (seconds to make a full scan). This can be completed without expensive or bulky capital equipment. Further, learning to interpret CRF measurements, while the physician does need to interpret, is straightforward and not nearly as subjective as interpreting confocal microscopy measurements, for example.

As a Surgical Tool for Assessing Clear Margins

The aim of surgical therapy for the treatment of lung cancer is to excise malignant tumors with clear margins meaning cancer cells are not present at the divided edge of the tissue that was removed [19]. The definition of an adequate margin width is loosely defined as 2 cm or the same as the diameter of the tumor. The tenet that a positive margin (cancer cells present at the cut edge of the excised tissue) results in an increased incidence of cancer recurrence is widely accepted. A “positive” margin is, in a sense, a surrogate for residual tumor, and those patients have been shown to have an increased risk of local recurrence and mortality [19-21].

Pathology evaluation of margins involves sampling the edges of tissue from all cut faces of the excised specimen. It’s impractical for a pathologist to microscopically examine the entire margin in a specimen. The technology presented allows for electrodes to be embedded on a handheld or robotic surgical tool allowing the surgeon to assess the margin quickly and effectively in the operating theater. In circumstances of a positive margin, adjuvant therapy can be implemented to help eliminate residual local disease [20]. Classification of tumors as microscopically positive (R1) versus tumors with margins grossly involved by tumor (R2) could be impacted by such a point-of-care analysis of a specimen [20,22-27].

As a Tool for Collecting Efficient Biopsies

Nondiagnostic rates for biopsies in CT-guided needle biopsies are reported to be 27.6%. Nodules presented smaller than 1 cm have a 40% nondiagnostic rate [28]. One out of 4 biopsies is inefficient, exposing the patient to complications and risk. The electrodes in the system demonstrated herein can be miniaturized and can fit within a 2 mm orifice of the bronchoscope or on the surface of a biopsy needle. Being able to have a device that can give an indication of malignancy without having to directly sample the tissue would give this process a much higher yield as some of those nondiagnostic samples could also come with a CRF that demonstrates the nodule is malignant. The electrodes can be spaced to assess the entirety of a nodule allowing to compensate for the heterogeneity of the nodule. The measurement can provide an accurate, real-time answer on the presence of cancer. Lesions that were previously difficult to locate from the airway in real time can be easily located as the technology can visualize macroscopically from the airway and identify both cancerous and benign nodules. This provides a precise periprocedural intralesional assessment mechanism to confirm the location of cancerous cells within a lesion.

Lymph Node Assessment

Lymph node status can often influence the decision to proceed with surgery first or induction chemotherapy (N2 nodes). Further, an N3 lymph node will influence the decision of whether or not surgical resection is indicated at all. These nodes can be assessed either by mediastinoscopy before a planned resection or during the planned resection but before the resection

itself has taken place. Real-time nodal assessment with the elimination of the time and resource-intensive frozen pathologic analytical process and its inherent uncertainties would be invaluable in these circumstances.

Residual Tumor

There are situations where a patient receives induction chemotherapy and/or radiation therapy followed by surgical resection. This involves 5% to 10% of lung cancer patients but a much larger proportion of esophageal cancer patients. Sometimes their preoperative induction therapy results in complete pathologic response (CPR), in which no viable tumor is detected via pathologic evaluation. Measuring a CRF in this patient population could potentially significantly impact this assessment. In the setting of a CPR, a positive CRF might put that designation into question and have some implications as to recurrence risk. Similarly, in patients who are treated with definitive chemoradiation therapy who are not surgical candidates and are under post-treatment surveillance, the ability to assess the treatment area with a CRF could potentially lead to earlier detection of recurrence. Furthermore, this could prove valuable if a local endoscopic resection margin is uncertain.

Conflicts of Interest

LB, OF, DC, AG, and MG are employed at NovaScan, Inc. NovaScan has patents pending on technology related to the research reported here. Advocate Aurora Health is a stockholder in NovaScan, but DD has no financial interest in the company. DD is also the robotic thoracic surgery proctor and speaker for Intuitive Surgical, Inc.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018 Nov 12;68(6):394-424 [FREE Full text] [doi: 10.3322/caac.21492] [Medline: 30207593]
2. John R, Ross H. The Global Economic Cost of Cancer. American Cancer Society and the LIVESTRONG Organization. 2010. URL: http://phrma-docs.phrma.org/sites/default/files/pdf/08-17-2010_economic_impact_study.pdf [accessed 2021-02-10]
3. Pulmonary Nodules - Health Encyclopedia - University of Rochester Medical Center. URL: https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=22&contentid=pulmonarynodules [accessed 2021-11-09]
4. Harzheim D, Eberhardt R, Hoffmann H, Herth FJ. The Solitary Pulmonary Nodule. Respiration 2015 Jun 26;90(2):160-172 [FREE Full text] [doi: 10.1159/000430996] [Medline: 26138915]
5. Bubendorf L, Lantuejoul S, de Langen AJ, Thunnissen E. Nonsmall cell lung carcinoma: diagnostic difficulties in small biopsies and cytological specimens: Number 2 in the Series "Pathology for the clinician" Edited by Peter Dorfmüller and Alberto Cavazza. Eur Respir Rev 2017 Jun 30;26(144):170007 [FREE Full text] [doi: 10.1183/16000617.0007-2017] [Medline: 28659503]
6. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. N Engl J Med 2020 Feb 06;382(6):503-513. [doi: 10.1056/NEJMoa1911793] [Medline: 31995683]
7. Shell J, Gregory WD. Efficient Cancer Detection Using Multiple Neural Networks. IEEE J Transl Eng Health Med 2017;5:2800607 [FREE Full text] [doi: 10.1109/JTEHM.2017.2757471] [Medline: 29282435]
8. Gregory WD, Marx JJ, Gregory CW, Mikkelson WM, Tjoe JA, Shell J. The Cole relaxation frequency as a parameter to identify cancer in breast tissue. Med Phys 2012 Jul 18;39(7):4167-4174. [doi: 10.1118/1.4725172] [Medline: 22830750]
9. Svoboda RM, Gharia MJ, Shell J, Gregory WD. Bioimpedance measurement as an assessment of margin positivity in Mohs surgical specimens of nonmelanoma skin cancer: Management implications. J Am Acad Dermatol 2018 Sep;79(3):591-593. [doi: 10.1016/j.jaad.2018.02.075] [Medline: 29544743]
10. McAdams ET, Jossinet J. Tissue impedance: a historical overview. Physiol Meas 1995 Aug 01;16(3 Suppl A):A1-13. [doi: 10.1088/0967-3334/16/3a/001] [Medline: 8528108]
11. Cole KS, Cole RH. Dispersion and Absorption in Dielectrics I. Alternating Current Characteristics. The Journal of Chemical Physics 1941 Apr;9(4):341-351. [doi: 10.1063/1.1750906]
12. Freeborn TJ, Maundy B, Elwakil A. Numerical extraction of Cole-Cole impedance parameters from step response. NOLTA 2011;2(4):548-561. [doi: 10.1587/nolta.2.548]
13. Freeborn T, Fu B. Fatigue-Induced Cole Electrical Impedance Model Changes of Biceps Tissue Bioimpedance. Fractal Fract 2018 Oct 24;2(4):27. [doi: 10.3390/fractalfract2040027]
14. Tang C, You F, Cheng G, Gao D, Fu F, Dong X. Modeling the frequency dependence of the electrical properties of the live human skull. Physiol Meas 2009 Dec 20;30(12):1293-1301. [doi: 10.1088/0967-3334/30/12/001] [Medline: 19843982]
15. Schwan H. Electrical properties of tissue and cell suspensions. In: Adv Biol Med Phys. In: Advances in Biological and Medical Physics; 1957:147-209.
16. SignaGel Electrode Gel: Parker Laboratories, Inc. URL: https://www.parkerlabs.com/signagel.asp [accessed 2021-11-09]
17. de Sousa VML, Carvalho L. Heterogeneity in Lung Cancer. Pathobiology 2018 Apr 10;85(1-2):96-107 [FREE Full text] [doi: 10.1159/000487440] [Medline: 29635240]

18. Nakajima T, Anayama T, Matsuda Y, Hwang DM, McVeigh PZ, Wilson BC, et al. Orthotopic lung cancer murine model by nonoperative transbronchial approach. Ann Thorac Surg 2014 May;97(5):1771-1775. [doi: 10.1016/j.athoracsur.2014.01.048] [Medline: 24792261]
19. Owen RM, Force SD, Gal AA, Feingold PL, Pickens A, Miller DL, et al. Routine intraoperative frozen section analysis of bronchial margins is of limited utility in lung cancer resection. Ann Thorac Surg 2013 Jun;95(6):1859-65; discussion 1865. [doi: 10.1016/j.athoracsur.2012.12.016] [Medline: 23415238]
20. Predina JD, Keating J, Patel N, Nims S, Singhal S. Clinical implications of positive margins following non-small cell lung cancer surgery. J Surg Oncol 2016 Mar 30;113(3):264-269. [doi: 10.1002/jso.24130] [Medline: 26719121]
21. Hancock JG, Rosen JE, Antonicelli A, Moreno A, Kim AW, Detterbeck FC, et al. Impact of adjuvant treatment for microscopic residual disease after non-small cell lung cancer surgery. Ann Thorac Surg 2015 Feb;99(2):406-413. [doi: 10.1016/j.athoracsur.2014.09.033] [Medline: 25528723]
22. He P, Yao G, Guan Y, Lin Y, He J. Diagnosis of lung adenocarcinoma in situ and minimally invasive adenocarcinoma from intraoperative frozen sections: an analysis of 136 cases. J Clin Pathol 2016 Dec 12;69(12):1076-1080. [doi: 10.1136/jclinpath-2016-203619] [Medline: 27174927]
23. Emerson LL, Layfield LJ. Solitary peripheral pulmonary papilloma evaluation on frozen section: a potential pitfall for the pathologist. Pathol Res Pract 2012 Dec 15;208(12):726-729. [doi: 10.1016/j.prp.2012.09.007] [Medline: 23131661]
24. Marchevsky AM, Changsri C, Gupta I, Fuller C, Houck W, McKenna RJ. Frozen section diagnoses of small pulmonary nodules: accuracy and clinical implications. Ann Thorac Surg 2004 Nov;78(5):1755-1759. [doi: 10.1016/j.athoracsur.2004.05.003] [Medline: 15511468]
25. Gupta R, McKenna R, Marchevsky AM. Lessons learned from mistakes and deferrals in the frozen section diagnosis of bronchioloalveolar carcinoma and well-differentiated pulmonary adenocarcinoma: an evidence-based pathology approach. Am J Clin Pathol 2008 Jul;130(1):11-20; quiz 146. [doi: 10.1309/UUH2XHECKEVD45PF] [Medline: 18550465]
26. Sawabata N, Ohta M, Matsumura A, Nakagawa K, Hirano H, Maeda H, Thoracic Surgery Study Group of Osaka University. Optimal distance of malignant negative margin in excision of nonsmall cell lung cancer: a multicenter prospective study. Ann Thorac Surg 2004 Feb;77(2):415-420. [doi: 10.1016/S0003-4975(03)01511-X] [Medline: 14759408]
27. Gupta R, McKenna R, Marchevsky AM. Lessons learned from mistakes and deferrals in the frozen section diagnosis of bronchioloalveolar carcinoma and well-differentiated pulmonary adenocarcinoma: an evidence-based pathology approach. Am J Clin Pathol 2008 Jul;130(1):11-20; quiz 146. [doi: 10.1309/UUH2XHECKEVD45PF] [Medline: 18550465]
28. Elicker BM. How Should We Manage Nondiagnostic Lung Biopsies? Radiology 2019 Mar;290(3):824-825. [doi: 10.1148/radiol.2018182630] [Medline: 30561275]

Abbreviations

CPE: constant phase element
CPR: complete pathologic response
CRF: Cole relaxation frequency
CT: computed tomography
NSCLC: nonsmall cell lung carcinoma
TRL: Toxicology Research Laboratory
UIC: University of Illinois

Edited by A Mavragani; submitted 01.12.21; peer-reviewed by P Zarogoulidis; comments to author 05.01.22; revised version received 06.01.22; accepted 12.01.22; published 21.02.22.

Please cite as:
Bogdanowicz L, Fidaner O, Ceres D, Grycuk A, Guidetti M, Demos D
The Cole Relaxation Frequency as a Parameter to Identify Cancer in Lung Tissue: Preliminary Animal and Ex Vivo Patient Studies
JMIR Biomed Eng 2022;7(1):e35346
URL: https://biomedeng.jmir.org/2022/1/e35346
doi:10.2196/35346
PMID:

Original Paper

Equity-Driven Sensing System for Measuring Skin Tone–Calibrated Peripheral Blood Oxygen Saturation (OptoBeat): Development, Design, and Evaluation Study

Alexander T Adams[1,2], BSc, MSc, PhD; Ilan Mandel[1], BSc; Yixuan Gao[1], BSc; Bryan W Heckman[2,3,4], PhD; Rajalakshmi Nandakumar[1], PhD; Tanzeem Choudhury[1], PhD

[1]Information Science, Cornell Tech, New York, NY, United States

[2]The Center for the Study of Social Determinants of Health, Meharry Medical College, Nashville, TN, United States

[3]Psychiatry and Behavioral Sciences, School of Medicine, Meharry Medical College, Nashville, TN, United States

[4]Division of Public Health, School of Graduate Studies and Research, Meharry Medical College, Nashville, TN, United States

Corresponding Author:
Alexander T Adams, BSc, MSc, PhD
Information Science
Cornell Tech
2 West Loop Rd
New York, NY, 10044
United States
Phone: 1 646 971 3777
Email: ata56@cornell.edu

Abstract

Background: Many commodity pulse oximeters are insufficiently calibrated for patients with darker skin. We demonstrate a quantitative measurement of this disparity in peripheral blood oxygen saturation (SpO_2) with a controlled experiment. To mitigate this, we present OptoBeat, an ultra–low-cost smartphone-based optical sensing system that captures SpO_2 and heart rate while calibrating for differences in skin tone. Our sensing system can be constructed from commodity components and 3D-printed clips for approximately US $1. In our experiments, we demonstrate the efficacy of the OptoBeat system, which can measure SpO_2 within 1% of the ground truth in levels as low as 75%.

Objective: The objective of this work is to test the following hypotheses and implement an ultra–low-cost smartphone adapter to measure SpO_2: skin tone has a significant effect on pulse oximeter measurements (hypothesis 1), images of skin tone can be used to calibrate pulse oximeter error (hypothesis 2), and SpO_2 can be measured with a smartphone camera using the screen as a light source (hypothesis 3).

Methods: Synthetic skin with the same optical properties as human skin was used in ex vivo experiments. A skin tone scale was placed in images for calibration and ground truth. To achieve a wide range of SpO_2 for measurement, we reoxygenated sheep blood and pumped it through synthetic arteries. A custom optical system was connected from the smartphone screen (flashing red and blue) to the analyte and into the phone's camera for measurement.

Results: The 3 skin tones were accurately classified according to the Fitzpatrick scale as types 2, 3, and 5. Classification was performed using the Euclidean distance between the measured red, green, and blue values. Traditional pulse oximeter measurements (n=2000) showed significant differences between skin tones in both alternating current and direct current measurements using ANOVA (direct current: $F_{2,5997}=3.1170 \times 10^5$, $P<.01$; alternating current: $F_{2,5997}=8.07 \times 10^6$, $P<.01$). Continuous SpO_2 measurements (n=400; 10-second samples, 67 minutes total) from 95% to 75% were captured using OptoBeat in an ex vivo experiment. The accuracy was measured to be within 1% of the ground truth via quadratic support vector machine regression and 10-fold cross-validation ($R^2=0.97$, root mean square error=0.7, mean square error=0.49, and mean absolute error=0.5). In the human-participant proof-of-concept experiment (N=3; samples=3 × N, duration=20-30 seconds per sample), SpO_2 measurements were accurate to within 0.5% of the ground truth, and pulse rate measurements were accurate to within 1.7% of the ground truth.

Conclusions: In this work, we demonstrate that skin tone has a significant effect on SpO_2 measurements and the design and evaluation of OptoBeat. The ultra-low-cost OptoBeat system enables smartphones to classify skin tone for calibration, reliably measure SpO_2 as low as 75%, and normalize to avoid skin tone–based bias.

(JMIR Biomed Eng 2022;7(1):e34934) doi:10.2196/34934

KEYWORDS

mobile health; ubiquitous health; health equity; bias; pulse oximetry; oximetry; mHealth; health app; skin tone; oximeter; smartphone; sensor; heart rate; oxygen level; oxygen saturation; pulse; mobile phone

Introduction

Background

The measurement of blood oxygenation is necessary in acute medical emergencies and useful for tracking physical fitness [1] and shortening recovery times [2]. Blood oxygen saturation is an essential physiological signal that can be measured noninvasively optically or invasively via blood gas measurements. Peripheral blood oxygen saturation (SpO_2) measures the ratio of oxygen-saturated hemoglobin in the blood to hemoglobin not saturated with oxygen. Oxygen saturation levels <90% can result in a variety of symptoms and adverse health conditions.

Researchers have persistently documented how common pulse oximeters overestimate blood oxygen levels in patients with darker skin tones [3-6]. Commercial devices consistently overestimate oxygen saturation by as much as 2%, which could result in mistreated hypoxia or inadequate treatment. This paper focuses on the development of an ultra–low-cost, novel method for the measurement of SpO_2 that can be calibrated according to an individual's skin tone using a smartphone.

The COVID-19 pandemic has resulted in acute global shortages of necessary medical supplies, including pulse oximeters [7]. The "Silent Hypoxia" of patients with COVID-19 [8] increased the necessity for regular screening of blood oxygen levels, thus greatly increasing the demand for cheap and easily manufacturable pulse oximeters. Disruptions in supply chains were persistent [9], particularly in the silicon and chip markets [10] necessary for the construction of many medical devices. Our method makes regular pulse oximetry screenings possible for billions of pre-existing smartphone owners at a fraction of the cost of consumer pulse oximeters. Using a phone's screen and camera to measure the divergent absorption rates of blue and red light in hemoglobin, we can measure blood oxygen levels without any additional electronics. The system can be made using approximately US $1 of commodity plastics and a plastic ear clip, which can be easily 3D-printed or cheaply manufactured at scale.

The proliferation of smartphones over the past decade provides a platform upon which a variety of health apps can be built. The combination of high-quality sensors, increasingly powerful mobile computation, and internet connectivity enables the creation of medical sensing systems that already live in billions of users' pockets. More than 1 million mobile health apps are available on major mobile platforms [11]. This work builds on the subset of methods that use mobile phones to cheaply monitor and measure cardiovascular health. Using the camera built into most smartphones, we can calibrate for variance in skin tone that is not sufficiently captured by commodity pulse oximeters.

In this paper, we introduce the equity-driven design of OptoBeat, a novel, ultra–low-cost smartphone-based pulse oximeter. OptoBeat can determine a coefficient to normalize pulse oximetry readings for differences in skin tone. OptoBeat can measure SpO_2 within 1% accuracy of the gold standard. We validated OptoBeat in blood oxygen levels that ranged from healthy (95%-100%) to critical (as low as 75%), corresponding to hypoxia. In this paper, we prove the following three hypotheses: (1) skin tone has a significant effect on pulse oximeter measurements (hypothesis 1), (2) images of skin tone can be used to calibrate for pulse oximeter error (hypothesis 2), and (3) SpO_2 can be measured with a smartphone camera using the screen as a light source (hypothesis 3).

Through quantitative analysis, we demonstrate how skin tone affects pulse oximeters. We show that the unilateral effect of skin tone not only decreases the signal to noise ratio but also affects the ratio of ratios between the 2 sources, which is used to calculate SpO_2. We further demonstrate how this can be done with a smartphone and describe the design, development, and testing of an ultra–low-cost smartphone-based pulse oximeter. OptoBeat enables blood oxygen saturation monitoring by augmenting the smartphone's camera system, focusing the light source, and leveraging extant computing capacity. With this system, we can use a skin tone measurement to adjust the ratios of the 2 source frequencies transmitted through the skin to calibrate the measurement of blood oxygen for a more accurate reading.

Specifically, we present the following contributions: (1) an experiment and quantitative analysis of how skin tone affects traditional pulse oximeters and a demonstration of how this can be remedied; (2) the design of the OptoBeat optical sensing system and an experiment to validate the theory behind our design; and (3) the design, execution, and results from an ex vivo experiment that validates the accuracy of pulse oximetry from healthy and critical SpO_2 levels against the gold standard and the results of a human-participant experiment to validate the efficacy of the device.

Prior Work

Overview

Typically, pulse oximeters measure the ratio of oxygenated to deoxygenated hemoglobin through capillaries in the finger, as shown in Figure 1. To capture SpO_2, a pulse oximeter uses at

least two different electromagnetic wave sources at different wavelengths—one with a higher transmission in oxygenated hemoglobin and one with a higher transmission in deoxygenated hemoglobin, as shown in Figure 2. The ratio of the 2 wavelengths is used to determine the blood oxygen saturation [12]. Common sources used in commercial pulse oximeters are red light at approximately 660 nm and infrared (IR) light between 880 and 940 nm. This is due to both cost and the characteristics of the waves as larger wavelengths in the visible and near-IR spectrum transmit better through the skin [13].

Figure 1. Traditional pulse oximeter design.

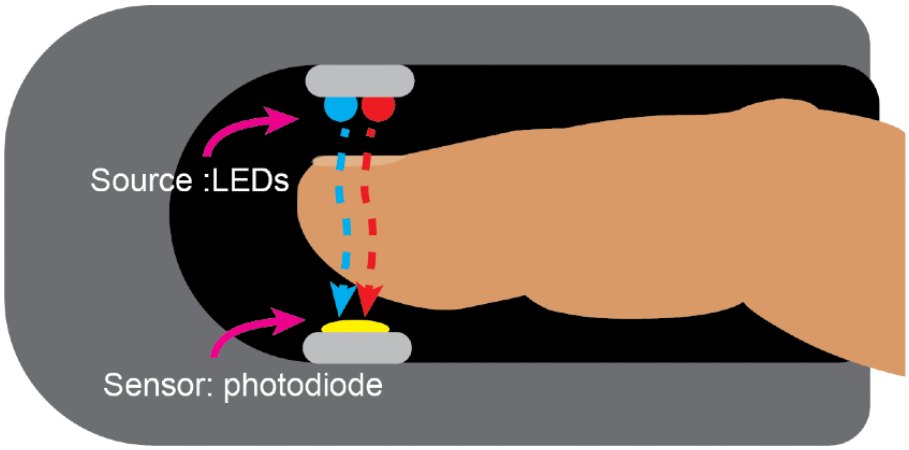
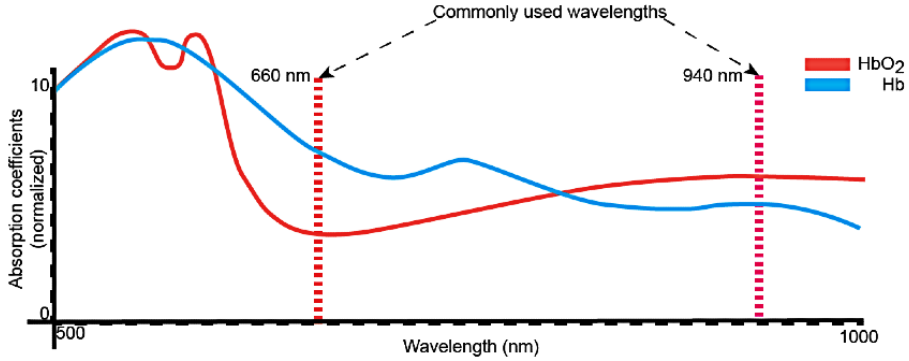


Figure 2. Absorption spectrum of oxygenated and deoxygenated hemoglobin and common wavelengths used in pulse oximetry. Hb: hemoglobin; HbO_2: oxygenated hemoglobin.



Pulse Oximetry and Skin Tone

Pulse oximetry measurements using standard commercial hardware can present biases and errors. Signal quality has been shown to decrease and increase with temperature when using finger-based measurements [14]. In the study by Nowara et al [15], the authors analyzed >400 videos of 73 individuals and found significant differences in the signal to noise ratio of those with darker skin tones. Sex is also predictive of errors in pulse oxygen measurements, resulting in higher SpO_2 estimates in women [16].

At saturation >80%, errors because of skin tone were not found to be significant; however, "in individuals with darkly pigmented skin, bias of up to 8% was observed at lower saturations" [17]. Anecdotal data showing this disparity have been documented as early as the late 1980s [17-19]. Analyses performed on patients from 2014 to 2015 and in 2020 demonstrated that "Black patients had nearly three times the frequency of acute hypoxemia that was not detected by pulse oximetry as White patients" [20]. Pulse oximeters are calibrated using correlation factors obtained by comparing the device being tested with a direct measurement of arterial hemoglobin saturation. The systematic absence of darker skinned participants during calibration may have resulted in devices optimized for whiter skin tones. Authors of a study stated the following:

> *In our 20 yr [sic] of testing pulse oximeter accuracy, and probably in other testing laboratories, the majority of subjects have been light skinned. Most pulse oximeters have probably been calibrated using light-skinned individuals, with the assumption that skin pigment does not matter.[16]*

This is similarly backed up by the 2013 Food and Drug Administration guidance on device calibration, which states that the "study should have subjects with a range of skin pigmentation, including at least 2 darkly pigmented subjects or 15% of your subject pool" [21]. Calibrating devices using only 15% of participants with darker skin is likely an underestimation of the percentage of Americans with higher levels of skin pigmentation [22]. The failures of federal standards and their resultant measurement errors fit within the well-documented ways in which racism is encoded in medical science, research, and treatment [23-25]. The disparities in measurement have become more pressing in the face of COVID-19, a disease in which morbidity disproportionately affects Black people [26] and exhibits an unusual form of silent hypoxemia [27].

Skin tone presents common complications in both dermal and transdermal optical measurements, particularly those in the visible spectrum. To help with this issue, researchers have found it useful to have a reference. One such standard reference is the Fitzpatrick skin tone scale. Although it is not without its own racial limitations [28], we found that it provided sufficient variation for understanding how pulse oximeters can fail in patients with darker skin tones.

Previous work has demonstrated how the Fitzpatrick scale was used to calibrate video-based heart rate monitoring in people with acute hypoxia [1]. Capturing individual differences in skin tone allows for the calibration of measurement bias in ways that are not possible with traditional pulse oximeters.

Smartphone Pulse Oximetry

The rapid development of smartphone sensors has enabled the design and implementation of smartphone-based medical and health-sensing systems. Systems that measure cardiac signals and blood composition, including heart rate [29,30], electrocardiogram [31,32], and hemoglobin levels [33], have been developed and deployed.

Various apps have been developed in the past decade for mobile phone–based pulse oximeters [34-37]. Smartphone-based pulse oximeters typically use the phone's flash as a light source and the rear-facing camera as a receiver. The camera has a red, green, and blue (RGB) filter over each pixel; the captured photoplethysmogram is broken down into blue and red light and then analyzed for SpO_2 [37]. However, the phone's flash is designed to transmit bursts of bright white light and is not stable, which would require a constant current source driver circuit for the flash's light-emitting diode (LED). There is a substantial imbalance between red and blue light that must be accounted for. Instability may arise from small shifts in the spectral emission pattern of the LED, the wide-range wavelength captured by the blue filter, and the flash's power supply [38]. In healthy patients, the performance of these systems [39] is in line with that of clinical pulse oximeters [40]. However, these systems perform poorly in unhealthy patients, particularly when their SpO_2 drops below 95% [40].

Commercial smartphone apps lack detailed information on the underlying algorithms and analyses used [3,41]. Furthermore, these systems are nonconcordant with ground truth measurements and regularly fail to detect hypoxia [42]. Researchers have experimented with modifying smartphones with optical filters as well, which has shown better results but still does not have sufficient accuracy outside of healthy SpO_2 levels [43].

Alternative smartphone-based systems incorporate an external sensor using the phone as a data hub. Many are standard fingertip-style pulse oximeters [36] with additional communication hardware such as Bluetooth Low Energy or Wi-Fi. Other smartphone-enabled form factors that are commercially available are wrist-worn and ring-based pulse oximeters [44]. Supplemental hardware can increase the accuracy of SpO_2 measurements with costs ranging from US $30 to US $400.

Theory: Blue Light Pulse Oximetry

We validated the electromagnetic spectrum emitted by the screen (white, all LEDs at full brightness) using a spectrometer [45]. The center (peak) band of the RGB LEDs in the screen for blue (465 nm) provided the required inverse relationship to hemoglobin that the center band of red (620 nm) has. Theoretically, this means that we can measure blood oxygen saturation with the LEDs built into a phone's screen.

Methods

Overview

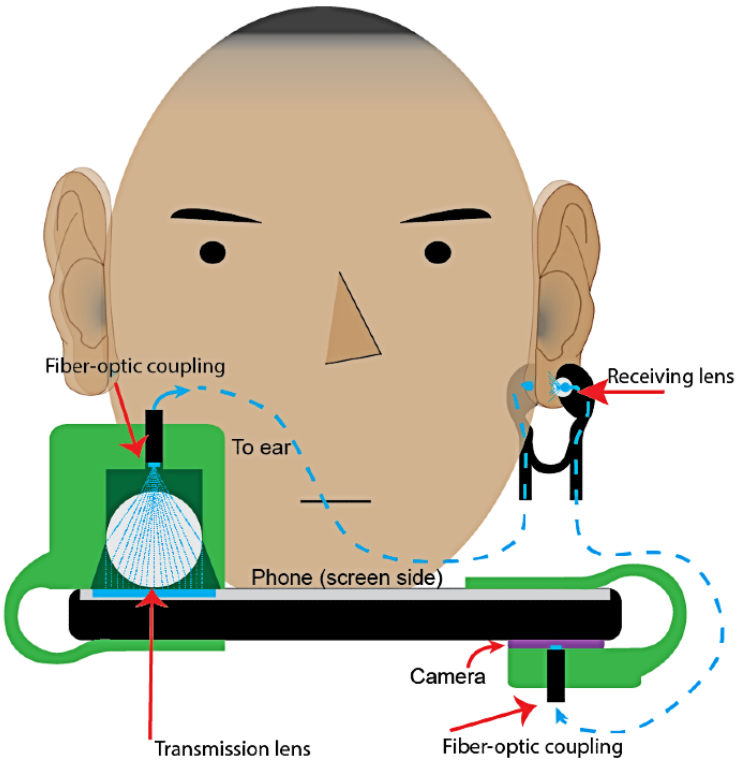
The OptoBeat system has several components to it, each of which we tested individually in a series of ex vivo experiments. These experiments allowed for control that was otherwise extremely difficult or impossible in a human-participant trial. For example, it would be extremely difficult to isolate skin from other components of the body (ie, blood, bone, and fat) in vivo. It is also dangerous for a healthy individual's blood oxygen saturation to fall below 95%. To test the range of oxygen saturation required to satisfy the needs of the OptoBeat system, we had to control the full range of blood oxygen saturation required for medical diagnosis. In this section, we describe these experiments in detail as well as the design of the OptoBeat system and how it was used to test our three hypotheses (hypotheses 1-3).

The OptoBeat Optical System

The final design consists of three 3D-printed plastic clips, two 3-foot fiber-optic cables, 1 acrylic ball lens, and some rubber bands or O-rings for grip. The total cost is <US $1 to put together individually; if scaled up, it would likely drop in price significantly. Figure 3 shows a cross-sectional diagram of how it all fits together with the smartphone.

Our goal was to focus as much light as possible and then couple the system with fiber-optic cables to maintain signal strength and block ambient light from entering the source (screen) or the receiver (camera). We evaluated dozens of commercial optical tools and systems to obtain an idea of what had been done; how we could move light around; and what materials were readily available, machinable, moldable, and affordable. This included optical filters, gratings, mirrors, and fiber plates, among others, that ranged from US $100 for individual components to tens of thousands of dollars.

Figure 3. Cross-sectional diagram of the OptoBeat system.



Fabrication

We limited our design exploration to materials and fabrication techniques such that the system could be built without specialized equipment. Most of the design, testing, and fabrication was executed using a laser cutter, a 3D printer, diamond files, and a drill press. In addition, to make an affordable, robust optical system, we explored a variety of raw materials. We found that the optimal shape for capturing light and coupling it with the fiber-optic cable was a long cone with a convex lens at the base. This design both collimated the beam and acted as a wave guide for coupling the light.

To make the lens, we experimented with hand-cutting a cast acrylic rod with diamond files and then polishing them to reach optical clarity. We also made a lens with a stereolithography 3D printer using clear resin that was then dipped in resin and cured to fill layer-height artifacts. In addition, we cast the negative of a lens (using a variety of materials) and then molded it with an optically clear urethane resin. Lens construction methods have proven to be labor-intensive, delicate tasks. Looking for alternative solutions to handmade or custom lenses, we found that a spherical lens worked nearly as well. This was mostly because focal length was not of concern as the distance to the fiber-optic cable could be easily adjusted. Cheap clear acrylic spheres are easily accessible from a variety of vendors and work well enough to focus light for transmissive pulse oximeters.

Placement

Given the limited brightness of the screen, we looked for other areas of the body in addition to the fingertip to capture SpO_2. Our search focused on areas that were physically accessible, with a short optical path (distance between transmitter and receiver) and a high density of capillaries to improve the pulsatile signal. Three plausible solutions were the fingertip (same as traditional pulse oximeters), the webbing between the fingers, and the earlobe. Data were captured at each location, with the earlobe showing the highest signal to noise ratio for output magnitude and pulsatile signal.

The earlobe is comfortable, and a sensor could easily be worn for prolonged periods without impeding the user in most activities (no more than a set of wired headphones would), affording more continuous monitoring.

Mobile App

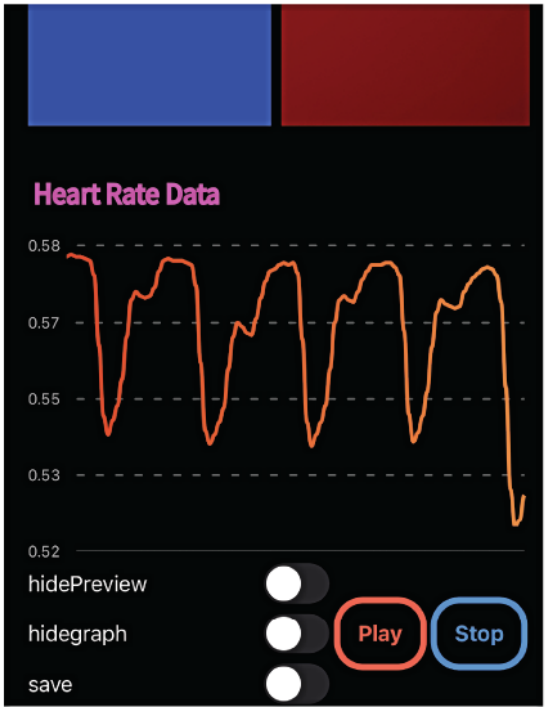
The OptoBeat signal acquisition app was designed on the iOS platform using Swift. The data acquisition page is shown in Figure 4. The app consists of the following components: data labels, camera selection, frame rate, sensitivity (ISO), exposure time, data visualization, pulse rate, and color selection. The app captures the light that has been emitted from the screen and passed through the user's skin. The mean values for RGB are stored on the phone for postanalysis.

Synchronizing the screen refresh rate and camera capture rate is crucial. The app ensures that the frequencies of data capture (camera) and source transmission (screen flashing) are synchronized. Misalignment results in the loss of essential information in the signal. An example of synchronization issues that we found is mixed-pulse capture. This occurs when half of the frame captures one color and the second half captures another. In our case, this was blue and red, resulting in a purple frame, which is unusable.

OptoBeat controls the camera using the Apple AVFoundation application programming interface. All automatic camera options are turned off, including autofocus, automatic white

balance, and low light boost. The app allows the experimenter to control all other settings, including ISO, exposure time, and camera capture rate. The camera runs on a separate thread, and each frame of the video is sent immediately to the camera buffer. Data are directly accessed through the CMSampleBuffer (Apple Inc), and the RGB values are logged to the back end using a data management pipeline in a new thread without slowing down the main process.

Figure 4. The OptoBeat mobile app.



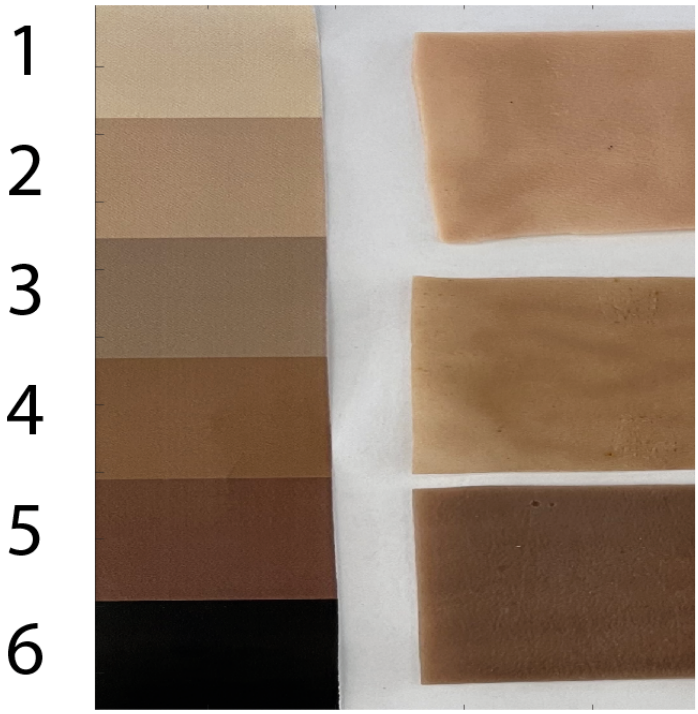
Skin Tone Experiment

As the behavior of electromagnetic waves in different skin tones is not necessarily continuous, we decided to classify instead of regress across the spectrum of skin. To demonstrate hypothesis 1 and test hypothesis 2, we classified the different tones as 1-6 according to the Fitzpatrick scale. For a ground truth, we printed a copy of the scale to use as a reference and placed it beside the synthetic skin [46] being tested (as shown in Figure 5) to account for variation in ambient lighting. Similar techniques have been used to measure bilirubin in the skin tone of infants to detect jaundice [47]. It should also be noted that synthetic skin is only available in 3 different tones; otherwise, we would have tested on more types.

The ground truth was determined by capturing the mean RGB value of each reference color. For each synthetic skin sample, the mean RGB was calculated and iteratively compared with the 6 ground truth classes using Euclidean distance and the absolute difference in luminance-weighted grayscale (Multimedia Appendix 1).

The results were identical for both grayscale and RGB; an example image is shown in Figure 5. Skin samples were matched to their corresponding groups on the Fitzpatrick scale in several images. The results were confirmed through visual inspection. The lighter sample matched group 2, the tan sample matched group 3, and the darker sample matched group 5.

Figure 5. Image of skin tone calibration strip and synthetic skin.

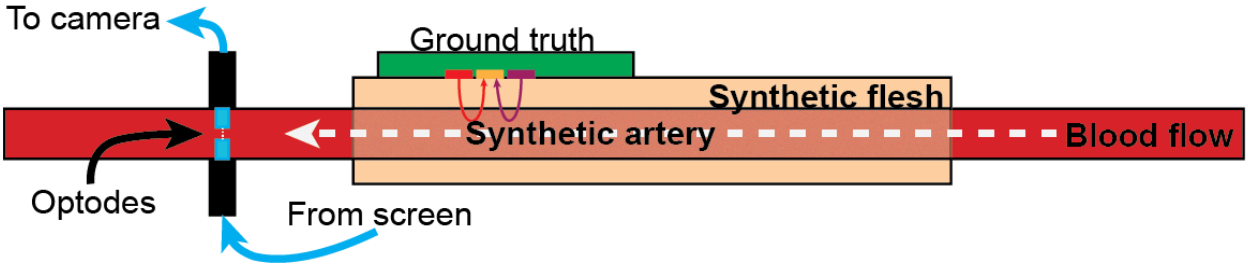


Blood Oxygen Saturation Experiment

After validating that we had obtained an optical signal from the earlobe through the video recording analysis of our system, we designed an experiment to test hypothesis 3, validating OptoBeat's SpO_2 measurements in mammalian blood using blue light pulse oximetry. In this experiment, we validated that, as hemoglobin deoxygenates, the transmission of blue light increases and that of red light decreases. This is based on the well-documented absorption and transmission spectrum of hemoglobin [12,48]. We designed this ex vivo experiment to allow us to control blood oxygen saturation over time and capture measurements that would be unhealthy, if not fatal, in human participants. Not only did this experiment allow us to obtain low blood oxygen saturation measurements, it also allowed us to capture continuous change over time, which is not feasible in human-participant trials. In this experiment, laboratory-grade sheep blood was oxygenated by pressurizing it in an oxygen-rich atmosphere. In the human lung, the P_{O2} "increases from approximately 40 mm Hg to 110 mm Hg, a pressure sufficient to ensure at least 95% saturation of hemoglobin with oxygen" [49]. The pressure tank was pressurized to 4137 mmHg (80 pounds of force per square inch) while a peristaltic pump connected to soapstone with a high density of microscopic pores circulated and diffused oxygen-rich air into the blood. This allowed the hemoglobin to bond to the air as it would in the capillaries of the lungs.

This setup demonstrated that the change in signal measurements could be attributed to blood oxygen saturation. After being oxygenated as in the previous experiment, sheep blood was pulsed through the artery at 60 beats per minute (bpm) to mimic a heart rate, and the blood was fed back into an open beaker, where it was exposed to standard air pressure and the hemoglobin could continue to deoxygenate. Figure 6 shows how the pulse oximeter was situated on the synthetic skin and artery to measure the ground truth. The experiment was run twice, collecting data continuously for several hours.

Figure 6. Cross-sectional diagram of the blood oxygen saturation experimental setup.



Human-Participant Proof of Concept

To validate that the OptoBeat system would work on human participants, data were collected on three of the authors. Data were collected 3 times for each author. The data collection period was approximately 30 seconds for each sample.

Ethics Approval

As there was no risk and the participants were all authors of this paper, we did not require Institutional Review Board approval. Data were collected between November 2020 and November 2021.

Results

The first experiments described in this section demonstrate the effect of skin tone on pulse oximeters. All the measurements in this first experiment are direct current (DC). These DC-component experiments demonstrate how these attributes affect the signal quality when isolated from the much more complex human body.

Skin Tone Experiment

We measured the transmission of red and IR light in the 3 different skin tones described in the *Methods* section. Experiments were performed using synthetic skin approximately 1 mm thick [46], which has the same physical and optical properties as human skin. By isolating the skin, we removed any potential confounding factors that might further influence the results, such as blood, SpO_2, skin thickness, or even hydration, among others.

By accessing the photodiode output of a commodity pulse oximeter, the values of the 660-nm (red) and 880-nm (IR) light as it passed through the different samples were recorded. It was important to record these values and not the resulting SpO_2 values as they had already been fitted to the model. To show that the problem lay in the actual hardware, we compared the red and IR values (Table 1), not the SpO_2 measurements, as this was where the calculations for SpO_2 were derived from. In addition, as we used distilled water to ensure that we were only showing the effect of the skin tone (as the SpO_2 would otherwise change over time), any SpO_2 measurement would be negligible. The ratios of the resulting 2 bands of light are shown in Table 1. We have also shown the percentage of change in ratios between the different skin tones calculated through a pairwise comparison of the different ratios.

For each skin tone, 2000 samples were collected, and the ratio of IR to red was calculated (type 5: SD 0.42%; type 3: SD 0.25%; type 2: SD 0.25%). We ran an ANOVA test and a pairwise comparison of the results. The results showed that the ratio of transmission for each skin tone was significantly different from the others ($F_{2,5997}=3.12 \times 10^5$; $P<.01$), which validated hypothesis 1.

To demonstrate hypothesis 2 (the use of skin tone data in calibrating pulse oximeters), the median sample, classified as type 3, was used as a reference to calibrate the others. For each skin tone, we derived coefficients to normalize the varying ratios between the different skin tones.

The ratio of the ratios was used to produce the following coefficients to normalize the skin tone absorbance of red and IR. The experiment was then replicated with our OptoBeat system using blue and red light, as shown in Table 2. The resulting ratios for hypothesis 2 were as follows (type 3 is always 1 as it was the reference skin tone): 0.8794 (red to IR) and 1.2396 (red to blue) for type 5 and 1.0326 (red to IR) and 1.0340 (red to blue) for type 2.

If it were the case that both wavelengths of light were affected equally, then the only errors that would arise would be due to changes in the signal to noise ratio. However, the change was unilateral, affecting the ratio of the wavelengths and not just the signal strength. Using the ratio of ratios between the alternating current (AC) and DC components of each wavelength did mitigate this to some extent. However, we hypothesized that the AC to DC ratio would also gain error across skin tones owing to their vastly different absorbency characteristics. To confirm this and further validate hypothesis 1, the same experiment as above was conducted using a pump system to move distilled water through a synthetic artery [46] and the 3 separate skin tones. Distilled water was used instead of blood to control for any adulterants, or actual changes in blood oxygen saturation.

As shown in Table 3, the results support the hypothesis that the AC to DC ratio of ratios (Equation 1) would also be affected by skin tone. The 2000 ratio-of-ratio samples showed little deviation from the mean (type 5: mean 0.77%, SD 0.11%; type 3: mean 0.81%, SD 0.11%; type 2: mean 0.94%, SD 0.22%).

Using an ANOVA test, the results were statistically significant ($F_{2,5997}=8.07 \times 10^6$; $P<.01$). Through pairwise multi-comparison, we see that the *R* value changes *by 4.3%* between skin types 5 and 3, *17.7%* between skin types 5 and 2, and *14%* between skin types 3 and 2. As predicted, the ratio of ratios using the AC signal did in fact lower the error; however, it was still significant, proving that, even with the ratio of ratios, hypothesis 1 still held true.

Table 1. Results from a direct current pulse oximeter skin tone experiment.

Skin tone type	Infrared value	Red value	Ratio of ratios	Difference from type 5, %	Difference from type 3, %	Difference from type 2, %
5	0.49	0.86	0.58	N/A[a]	12.1	14.8
3	0.51	1	0.51	12.1	N/A	3.2
2	0.50	0.96	0.52	14.8	3.2	N/A

[a]N/A: not applicable (data are compared to themselves).

Table 2. Results from a direct current OptoBeat skin tone experiment.

Skin tone type	Blue value	Red value	Ratio of ratios	Difference from type 5, %	Difference from type 3, %	Difference from type 2, %
5	0.8	0.61	0.76	N/A[a]	19.3	16.6
3	1	0.94	0.94	19.3	N/A	3.3
2	0.65	0.59	0.91	16.6	3.3	N/A

[a]N/A: not applicable (data are compared to themselves).

Table 3. Results from an alternating current pulse oximeter skin tone experiment.

Skin tone type	Ratio of ratios	Difference from type 5, %	Difference from type 3, %	Difference from type 2, %
5	0.77	N/A[a]	4.3	17.7
3	0.81	4.3	N/A	14
2	0.94	17.7	14	N/A

[a]N/A: not applicable (data are compared to themselves).

Blood Oxygen Saturation Experiment

The data (n=400; 10-second samples, 67 minutes total) were cleaned with a bandpass filter, with cutoff frequencies at 0.5 Hz and 2 Hz (30-120 bpm pass band), and then passed through a Savitzky-Golay filter for further smoothing (third order, 35-sample window). This provided a clear pulsatile signal, as shown in Figure 7. The AC, as measured by the root mean square, and DC (mean) components of the signal were extracted from each signal in 10-second windows (a common window size for this type of calculation) and then inserted into Equation 1 to calculate the ratio of ratios.

The resulting *R* value from Equation 1 has a direct relationship to the SpO_2 percentage and can be mapped using either a linear equation or regression. A linear mapping of the *R* values to SpO_2 ($f(x) = mx + b$; m=79.3, b=−5.7 with 95% confidence bounds) rendered very good results (R^2=0.94, root mean square error=1, mean square error=1.1, mean absolute error=0.89); however, we found that a quadratic support vector machine worked better with our data, as shown in Figure 8. This is most likely due to a change in the response curve when more blue light is present with lowering SpO_2 levels as the blue wavelength is much smaller and does not penetrate the skin as well as red or IR wavelengths. Each yellow dot in Figure 8 is an individual measurement. Each *R* value was calculated over 10 seconds of data. The orange line shows the error from the ground truth for each measurement. The performance, as shown in Figure 9, was validated with a 10-fold cross-validation (R^2=0.97, root mean square error=0.7, mean square error=0.49, mean absolute error=0.5). Figure 10 shows the residual error for each percentage of SpO_2 measured. The experiment showed that the OptoBeat system measured SpO_2 as low as *75%* within *–1%* to *+1%* of the ground truth, proving hypothesis 3.

OptoBeat measures continuously with a floating point, whereas the ground truth estimates integer values. We believe that the performance would be more strongly correlated if both were continuous. We plan to evaluate this in the future.

Figure 7. Pulsatile signal of red and blue lights captured in the oxygen saturation experiment.

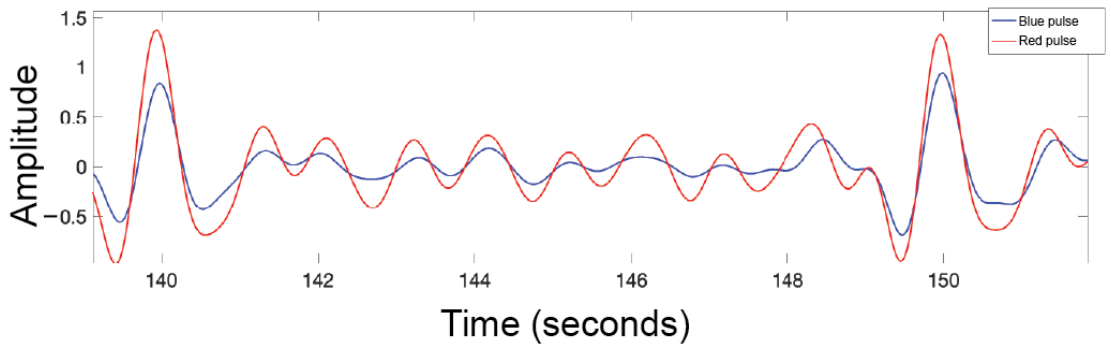


Figure 8. Results of the quadratic support vector machine (SVM) mapping R values captured from OptoBeat to the ground truth. SpO_2: peripheral blood oxygen saturation.

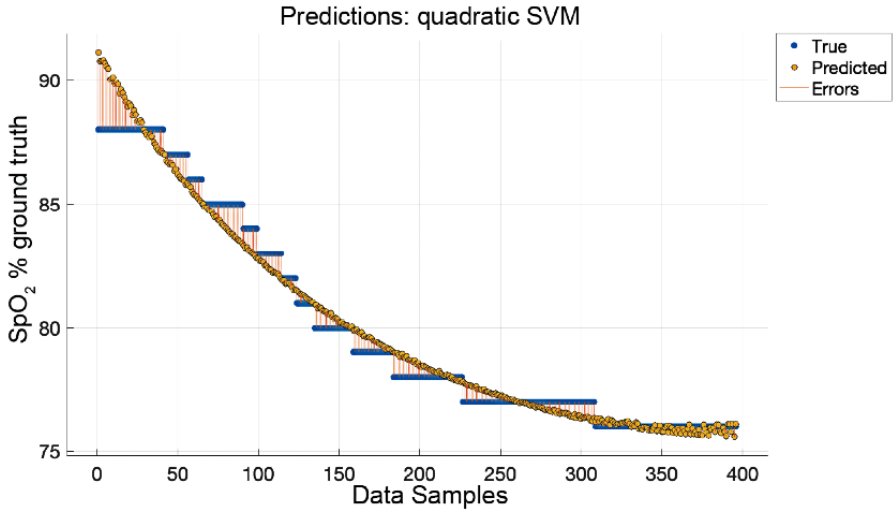


Figure 9. Predicted versus true response from regression.

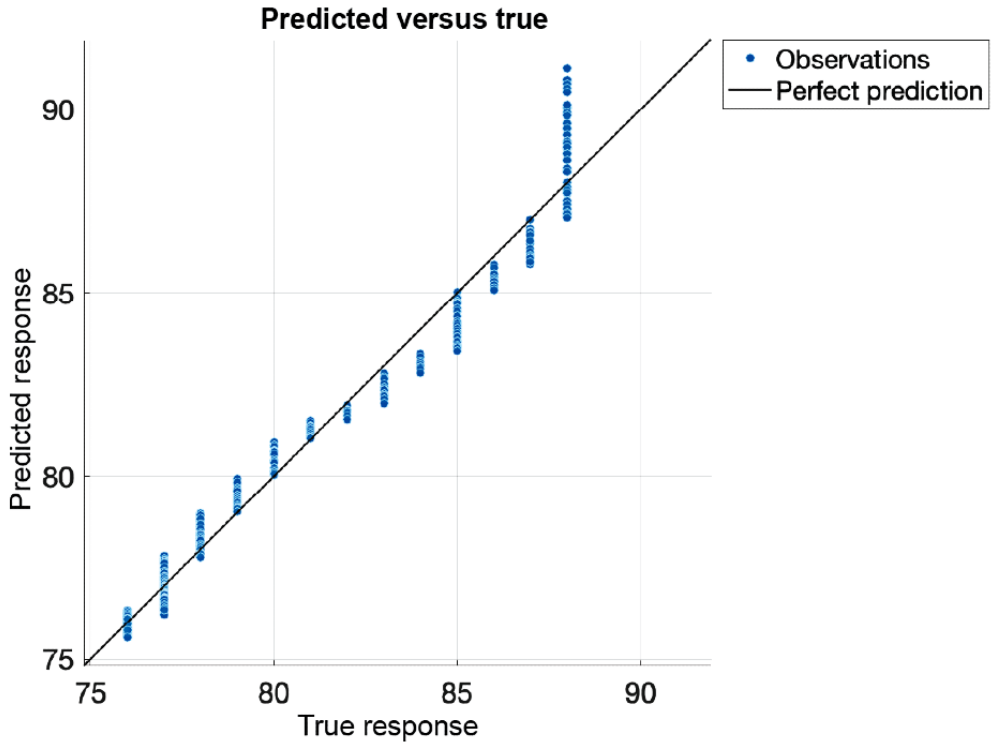
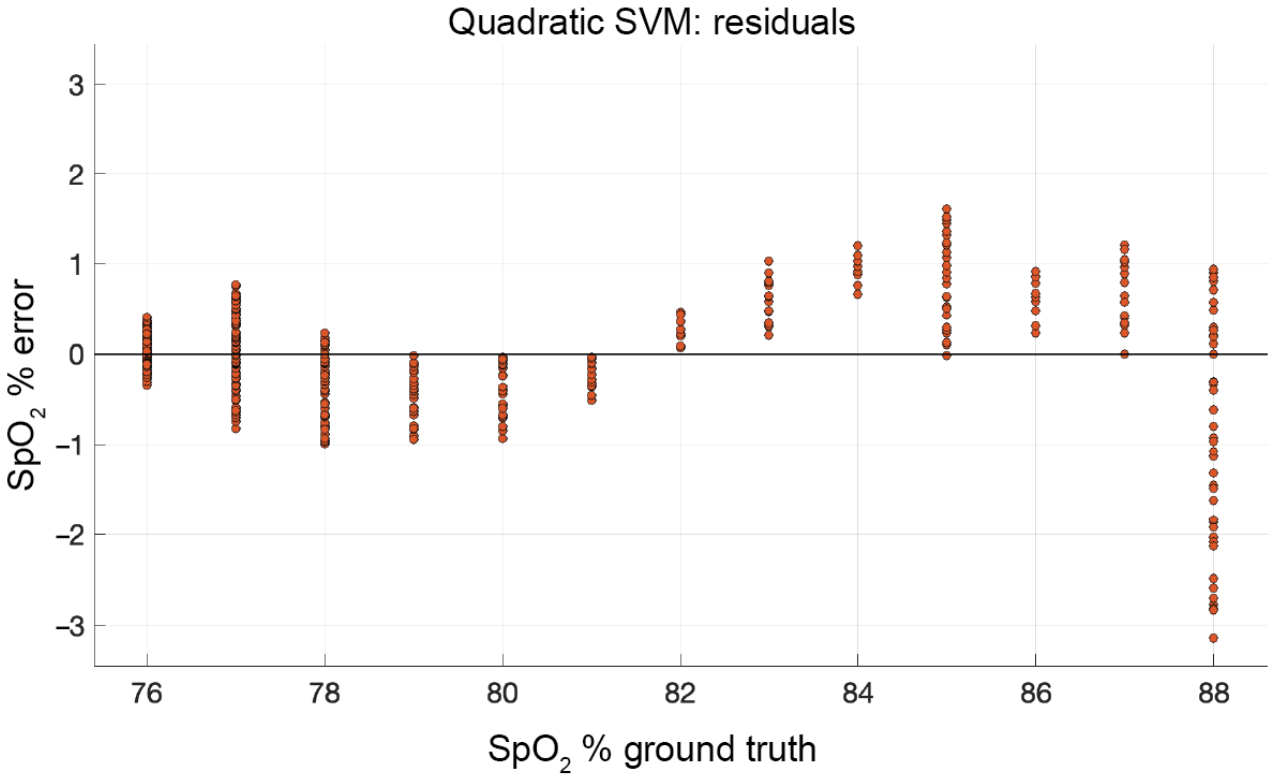


Figure 10. Residual error: OptoBeat and the ground truth. SpO_2: peripheral blood oxygen saturation; SVM: support vector machine.



Human-Participant Proof of Concept

Each of the participants (N=3; samples=3 × N, duration=20-30 seconds/sample) measured within –1% to +1% of the ground truth, a commodity pulse oximeter. Figure 11 shows the pulsatile signal captured by OptoBeat. In this sample (23 seconds) from 33% (1/3) of the participants, the heart rate measured by OptoBeat was approximately 111 bpm, and the ground truth was 114 bpm (error=1.7%).

Figure 12 shows the resulting SpO_2 measurement as captured by OptoBeat from the same data shown on the left. Both the polynomial trend and the raw, individual sample measurements are shown. It can be seen that the ground truth, which is only integer values, is *<0.5%* of the OptoBeat measurement. We believe, as the ground truth is rounded, that this corresponds to the spike (at approximately 17.5 seconds) in the OptoBeat measurement that exceeds *98.5%*. Although each person was healthy and the range of measured blood oxygen saturation was small (between *97*% and *99*% SpO_2), this demonstrates the feasibility of using OptoBeat in human participants.

Figure 11. Pulsatile signal of red and blue lights captured in the human-participant proof-of-concept experiment.

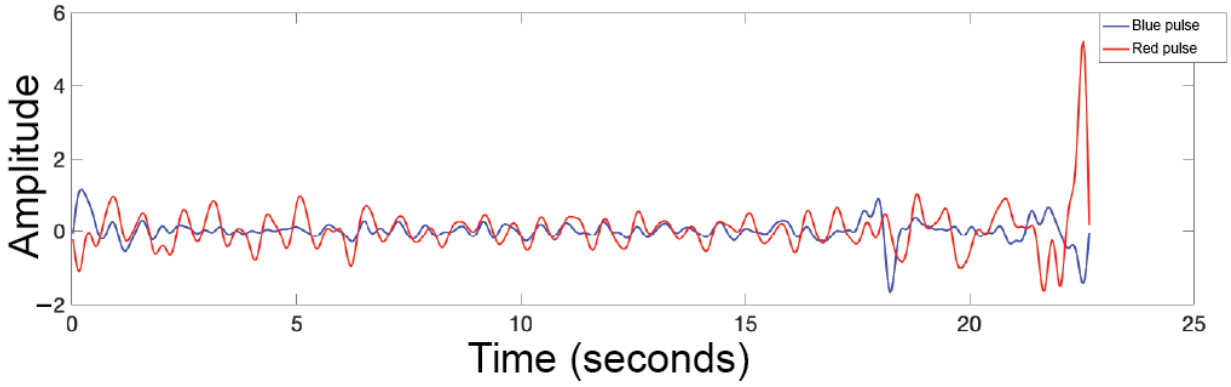
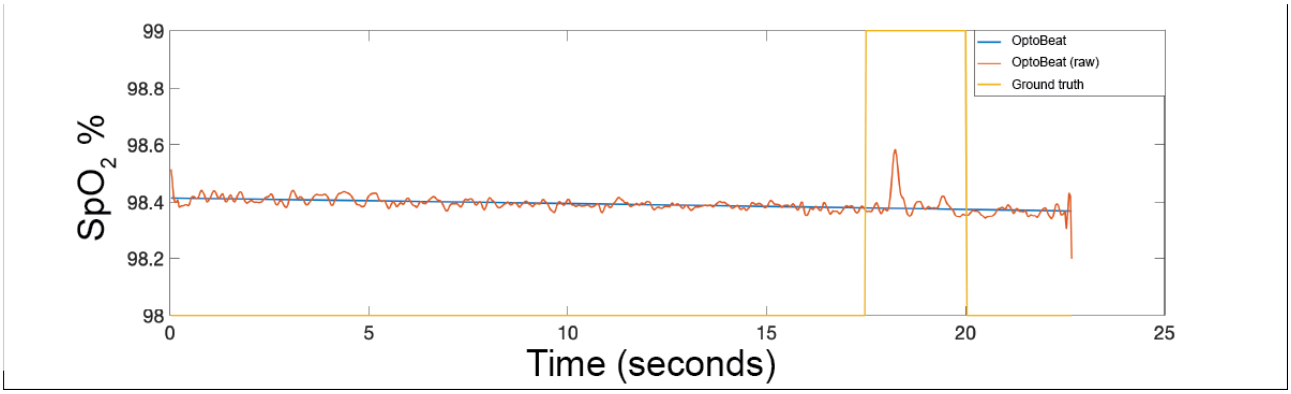


Figure 12. OptoBeat's peripheral blood oxygen saturation calculations against the ground truth in the human-participant proof-of-concept experiment.



Discussion

Principal Findings

In this paper, we have presented the design and evaluation of OptoBeat, an optical attachment for smartphones that can reliably measure SpO_2 and calibrate the measurements according to skin tone via images. The phone-based system for oxygen saturation measurement has potential benefits compared with existing pulse oximeters. With the added benefit of the smartphone's hardware and computing power, we can not only measure SpO_2 but also skin tone, which can be used to account for errors in SpO_2 measurements.

Strengths and Limitations

Our system is cheaper and simpler to produce than most commercial pulse oximeters. Manufactured at scale, it could be used to shift from a regime of rigid thresholds for admittance or treatment at the point of care to home measurement. At home, trending measures of pulse oximetry could be used to monitor clinical progress or detect silent hypoxemia early [50].

Although this work demonstrates the efficacy of OptoBeat in an ex vivo laboratory experiment, our human-participant proof of concept was only meant to show feasibility, not to validate it for human use. Clinical studies would have to be conducted on a large, diverse population to validate this.

Future Directions

Moving forward, we plan to deploy and test the current OptoBeat system in a clinical setting. We have partnered with our medical school to take measurements while patients are undergoing cardiothoracic surgery as this will give us access to a range of blood oxygen measurements while a ground truth is being collected without adding additional risk to participants. Furthermore, we plan to redesign traditional pulse oximeters using what we learned from OptoBeat to develop a stand-alone device that can account for variations in skin tone. This will include a large-scale skin tone data collection to build a model that can be leveraged by the device.

Clinical Implications and Conclusions

Wearable devices currently stand to exacerbate existing health disparities in underserved racial populations who would benefit most from enhanced detection and treatment of health issues. Indeed, such racial disparities occur in the context of more significant cultural issues and reflect the mistrust that many underserved racial populations have for the medical system. This mistrust is primarily due to the medical community's historical bias toward addressing White Americans' health needs at the expense of underserved racial populations' health and well-being. There is an urgent need to overcome the cycle of disparities produced by social injustice across conditions in the places where people live, learn, work, and play. Ensuring that wearables and remote patient monitoring tools are equally efficacious across different populations is necessary to accelerate health equity in the populations who would benefit most from such technology.

Acknowledgments

This project was supported by division of Information and Intelligent Systems (RAPID; award number 2031977) of the National Science Foundation. It was also supported by the National Institute on Drug Abuse (grant K23DA041616), the National Institute on Minority Health and Health Disparities (grant P50MD017347), and the National Institute of Allergy and Infectious Diseases (grant P30AI110527) of the National Institutes of Health.

Conflicts of Interest

TC is the cofounder of HealthRhythms, Inc, and also the Senior Vice President of Digital Health at the Optum Labs (part of United Health Group). The other authors have no conflicts to declare.

Multimedia Appendix 1
Skin tone calibration equations.
[DOCX File , 13 KB - biomedeng_v7i1e34934_app1.docx]

References

1. Addison PS, Jacquel D, Foo DM, Borg UR. Video-based heart rate monitoring across a range of skin pigmentations during an acute hypoxic challenge. J Clin Monit Comput 2018;32(5):871-880 [FREE Full text] [doi: 10.1007/s10877-017-0076-1] [Medline: 29124562]
2. Baxter C, Carroll JA, Keogh B, Vandelanotte C. Assessment of mobile health apps using built-in smartphone sensors for diagnosis and treatment: systematic survey of apps listed in international curated health app libraries. JMIR Mhealth Uhealth 2020;8(2):e16741 [FREE Full text] [doi: 10.2196/16741] [Medline: 32012102]
3. Berry DC, Seitz SR. Educating the educator: use of pulse oximetry in athletic training. Athl Train Educ J 2012;7(2):74-80 [FREE Full text] [doi: 10.4085/070274]
4. Emery JR. Skin pigmentation as an influence on the accuracy of pulse oximetry. J Perinatol 1987;7(4):329-330. [Medline: 2463349]
5. Paradies Y, Ben J, Denson N, Elias A, Priest N, Pieterse A, et al. Racism as a determinant of health: a systematic review and meta-analysis. PLoS One 2015;10(9):e0138511 [FREE Full text] [doi: 10.1371/journal.pone.0138511] [Medline: 26398658]
6. Pulse Oximeter Accuracy and Limitations: FDA Safety Communication. U.S. Food & Drug Administration. 2021. URL: https://tinyurl.com/yckcebcs [accessed 2021-11-30]
7. McMahon DE, Peters GA, Ivers LC, Freeman EE. Global resource shortages during COVID-19: bad news for low-income countries. PLoS Negl Trop Dis 2020;14(7):e0008412 [FREE Full text] [doi: 10.1371/journal.pntd.0008412] [Medline: 32628664]
8. Wilkerson RG, Adler JD, Shah NG, Brown R. Silent hypoxia: a harbinger of clinical deterioration in patients with COVID-19. Am J Emerg Med 2020;38(10):2243.e5-2243.e6 [FREE Full text] [doi: 10.1016/j.ajem.2020.05.044] [Medline: 32471783]
9. Mahajan K, Tomar S. Here today, gone tomorrow: COVID-19 and supply chain disruptions. Am J Agric Econ (forthcoming) 2020:1-39 [FREE Full text] [doi: 10.2139/ssrn.3596720]
10. Ivanov D. Supply chain viability and the COVID-19 pandemic: a conceptual and formal generalisation of four major adaptation strategies. Int J Prod Res 2021;59(12):3535-3552 [FREE Full text] [doi: 10.1080/00207543.2021.1890852]
11. Villarreal V, Berbey-Alvarez A. Evaluation of mHealth applications related to cardiovascular diseases: a systematic review. Acta Inform Med 2020;28(2):130-137 [FREE Full text] [doi: 10.5455/aim.2020.28.130-137] [Medline: 32742066]
12. Zijlstra WG, Buursma A, van Assendelft OW. Visible and near infrared absorption spectra of human and animal haemoglobin: determination and application. Boca Raton, FL: CRC Press; 2000.
13. Gardner CM. Transmission versus reflectance spectroscopy for quantitation. J Biomed Opt 2018;23(1):1-8 [FREE Full text] [doi: 10.1117/1.JBO.23.1.018001] [Medline: 29297210]
14. Khan M, Pretty CG, Amies AC, Elliott R, Chiew YS, Shaw GM, et al. Analysing the effects of cold, normal, and warm digits on transmittance pulse oximetry. Biomed Signal Process Control 2016;26:34-41 [FREE Full text] [doi: 10.1016/j.bspc.2015.12.006]
15. Nowara EM, McDuff D, Veeraragh A. A meta-analysis of the impact of skin type and gender on non-contact photoplethysmography measurements. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020 Presented at: CVPRW '20; June 14-19, 2020; Seattle, WA p. 1148-1155. [doi: 10.1109/cvprw50498.2020.00150]
16. Feiner JR, Severinghaus JW, Bickler PE. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. Anesth Analg 2007;105(6 Suppl):S18-S23. [doi: 10.1213/01.ane.0000285988.35174.d9] [Medline: 18048893]
17. Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. Anesthesiology 2005;102(4):715-719 [FREE Full text] [doi: 10.1097/00000542-200504000-00004] [Medline: 15791098]
18. Jubran A, Tobin MJ. Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients. Chest 1990;97(6):1420-1425. [doi: 10.1378/chest.97.6.1420] [Medline: 2347228]
19. Ries AL, Prewitt LM, Johnson JJ. Skin color and ear oximetry. Chest 1989;96(2):287-290. [doi: 10.1378/chest.96.2.287] [Medline: 2752811]
20. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. N Engl J Med 2020;383(25):2477-2478 [FREE Full text] [doi: 10.1056/NEJMc2029240] [Medline: 33326721]
21. Pulse oximeters - premarket notification submissions [510(k)s]: guidance for industry and food and drug administration staff. U.S. Food & Drug Administration. 2013. URL: https://tinyurl.com/52vwp4xa [accessed 2021-11-30]
22. Parra EJ, Kittles RA, Shriver MD. Implications of correlations between skin color and genetic ancestry for biomedical research. Nat Genet 2004;36(11 Suppl):S54-S60. [doi: 10.1038/ng1440] [Medline: 15508005]
23. Evans MK, Rosenbaum L, Malina D, Morrissey S, Rubin EJ. Diagnosing and treating systemic racism. N Engl J Med 2020;383(3):274-276. [doi: 10.1056/NEJMe2021693] [Medline: 32521155]
24. Colvonen PJ, DeYoung PN, Bosompra NO, Owens RL. Limiting racial disparities and bias for wearable devices in health science research. Sleep 2020;43(10):zsaa159 [FREE Full text] [doi: 10.1093/sleep/zsaa159] [Medline: 32893865]
25. Williams DR, Mohammed SA. Racism and health I: pathways and scientific evidence. Am Behav Sci 2013;57(8):1152-1173 [FREE Full text] [doi: 10.1177/0002764213487340] [Medline: 24347666]

26. Golestaneh L, Neugarten J, Fisher M, Billett HH, Gil MR, Johns T, et al. The association of race and COVID-19 mortality. EClinicalMedicine 2020;25:100455 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100455] [Medline: 32838233]
27. Tobin MJ, Laghi F, Jubran A. Why COVID-19 silent hypoxemia is baffling to physicians. Am J Respir Crit Care Med 2020;202(3):356-360 [FREE Full text] [doi: 10.1164/rccm.202006-2157CP] [Medline: 32539537]
28. Ware OR, Dawson JE, Shinohara MM, Taylor SC. Racial limitations of fitzpatrick skin type. Cutis 2020;105(2):77-80. [Medline: 32186531]
29. Gregoski MJ, Mueller M, Vertegel A, Shaporev A, Jackson BB, Frenzel RM, et al. Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. Int J Telemed Appl 2012;2012:696324 [FREE Full text] [doi: 10.1155/2012/696324] [Medline: 22272197]
30. Jonathan E, Leahy M. Investigating a smartphone imaging unit for photoplethysmography. Physiol Meas 2010;31(11):N79-N83. [doi: 10.1088/0967-3334/31/11/N01] [Medline: 20871134]
31. de Lucena SE, Sampaio DJ, Mall B, Meyer M, Burkart MA, Keller FV. ECG monitoring using Android mobile phone and Bluetooth. In: Proceedings of the IEEE International Instrumentation and Measurement Technology Conference. 2015 Presented at: I2MTC '15; May 11-14, 2015; Pisa, Italy p. 1976-1980 URL: https://doi.org/10.1109/i2mtc.2015.7151585 [doi: 10.1109/i2mtc.2015.7151585]
32. Wen C, Yeh MF, Chang KC, Lee RG. Real-time ECG telemonitoring system design with mobile phone platform. Measurement 2008;41(4):463-470 [FREE Full text] [doi: 10.1016/j.measurement.2006.12.006]
33. Wang EJ, Li W, Hawkins D, Gernsheimer T, Norby-Slycord C, Patel SN. HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016 Presented at: UbiComp '16; September 12-16, 2016; Heidelberg, Germany p. 593-604 URL: https://doi.org/10.1145/2971648.2971653 [doi: 10.1145/2971648.2971653]
34. Oximeter. digiDoc technologies. Egersund, Norway URL: http://www.digidoctech.no/ [accessed 2021-11-30]
35. AppAdvice. URL: https://appadvice.com/app/heart-rate-pulse-oximeter/597118857 [accessed 2021-11-30]
36. IOximeter. safe heart. URL: http://safeheartus.com/ioximeter [accessed 2021-11-30]
37. Scully CG, Lee J, Meyer J, Gorbach AM, Granquist-Fraser D, Mendelson Y, et al. Physiological parameter monitoring from optical recordings with a mobile phone. IEEE Trans Biomed Eng 2012;59(2):303-306 [FREE Full text] [doi: 10.1109/TBME.2011.2163157] [Medline: 21803676]
38. Karlen W, Lim J, Ansermino JM, Dumont G, Scheffer C. Design challenges for camera oximetry on a mobile phone. Annu Int Conf IEEE Eng Med Biol Soc 2012;2012:2448-2451. [doi: 10.1109/EMBC.2012.6346459] [Medline: 23366420]
39. Tomlinson S, Behrmann S, Cranford J, Louie M, Hashikawa A. Accuracy of smartphone-based pulse oximetry compared with hospital-grade pulse oximetry in healthy children. Telemed J E Health 2018;24(7):527-535. [doi: 10.1089/tmj.2017.0166] [Medline: 29215972]
40. Tremper KK. Pulse oximetry. Chest 1989;95(4):713-715. [doi: 10.1378/chest.95.4.713] [Medline: 2647421]
41. Marek W, Marek E, Friz Y, Vogel P, Mückenhoff K, Kotschy-Lang N. A new procedure for the estimation of physical fitness of patients during clinical rehabilitation using the 6-minutes walk test. Pneumologie 2010;64(3):155-162 [FREE Full text] [doi: 10.1055/s-0029-1215233] [Medline: 20072958]
42. Jordan TB, Meyers CL, Schrading WA, Donnelly JP. The utility of iPhone oximetry apps: a comparison with standard pulse oximetry measurement in the emergency department. Am J Emerg Med 2020;38(5):925-928 [FREE Full text] [doi: 10.1016/j.ajem.2019.07.020] [Medline: 31471076]
43. Bui N, Nguyen A, Nguyen P, Truong H, Ashok A, Dinh T, et al. Smartphone-based SpO2 measurement by exploiting wavelengths separation and chromophore compensation. ACM Trans Sen Netw 2020;16(1):1-30. [doi: 10.1145/3360725]
44. Wellue O2Ring. Wellue Health. 2021. URL: https://getwellue.com/pages/o2ring-oxygen-monitor [accessed 2021-11-30]
45. nanoLambda. 2021. URL: https://nanolambda.myshopify.com/ [accessed 2021-11-30]
46. Syndaver. 2020. URL: https://syndaver.com [accessed 2021-11-30]
47. Mariakakis A, Banks MA, Phillipi L, Yu L, Taylor J, Patel SN. BiliScreen: smartphone-based scleral jaundice monitoring for liver and pancreatic disorders. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017;1(2):1-26 [FREE Full text] [doi: 10.1145/3090085]
48. Liu P, Zhu Z, Zeng C, Nie G. Specific absorption spectra of hemoglobin at different PO2 levels: potential noninvasive method to detect PO2 in tissues. J Biomed Opt 2012;17(12):125002 [FREE Full text] [doi: 10.1117/1.JBO.17.12.125002] [Medline: 23208210]
49. McNamara PJ, El-Khuffash A. Oxygen transport and delivery. In: Polin RA, Abman SH, Rowitch DH, Benitz WE, Fox WW, editors. Fetal and neonatal physiology. 5th edition. Amsterdam, The Netherlands: Elsevier; 2017:724-37.e2.
50. Teo J. Early detection of silent hypoxia in Covid-19 pneumonia using smartphone pulse oximetry. J Med Syst 2020;44(8):134 [FREE Full text] [doi: 10.1007/s10916-020-01587-6] [Medline: 32562006]

Abbreviations

AC: alternating current
bpm: beats per minute

DC: direct current
IR: infrared
LED: light-emitting diode
RGB: red, green, and blue
SpO_2: peripheral blood oxygen saturation

Edited by T Leung; submitted 13.11.21; peer-reviewed by A Mariakakis, K Ho; comments to author 30.01.22; revised version received 20.02.22; accepted 07.03.22; published 22.04.22.

Please cite as:
Adams AT, Mandel I, Gao Y, Heckman BW, Nandakumar R, Choudhury T
Equity-Driven Sensing System for Measuring Skin Tone–Calibrated Peripheral Blood Oxygen Saturation (OptoBeat): Development, Design, and Evaluation Study
JMIR Biomed Eng 2022;7(1):e34934
URL: https://biomedeng.jmir.org/2022/1/e34934
doi:10.2196/34934
PMID:38875699

Original Paper

A Novel Framework for Mixed Reality–Based Control of Collaborative Robot: Development Study

Md Tanzil Shahria[1*], BSc; Md Samiul Haque Sunny[1*], BSc; Md Ishrak Islam Zarif[2*], BSc; Md Mahafuzur Rahaman Khan[3*], BSc; Preet Parag Modi[3], BE; Sheikh Iqbal Ahamed[2*], PhD; Mohammad H Rahman[3*], PhD

[1]Department of Computer Science, University of Wisconsin-Milwaukee, Milwaukee, WI, United States

[2]Department of Computer Science, Marquette University, Milwaukee, WI, United States

[3]Department of Mechanical Engineering, University of Wisconsin-Milwaukee, Milwaukee, WI, United States

[*]these authors contributed equally

Corresponding Author:
Md Tanzil Shahria, BSc
Department of Computer Science
University of Wisconsin-Milwaukee
3200 N Cramer Street
Milwaukee, WI, 53211
United States
Phone: 1 4147376701
Email: mshahria@uwm.edu

Abstract

Background: Applications of robotics in daily life are becoming essential by creating new possibilities in different fields, especially in the collaborative environment. The potentials of collaborative robots are tremendous as they can work in the same workspace as humans. A framework employing a top-notch technology for collaborative robots will surely be worthwhile for further research.

Objective: This study aims to present the development of a novel framework for the collaborative robot using mixed reality.

Methods: The framework uses Unity and Unity Hub as a cross-platform gaming engine and project management tool to design the mixed reality interface and digital twin. It also uses the Windows Mixed Reality platform to show digital materials on holographic display and the Azure mixed reality services to capture and expose digital information. Eventually, it uses a holographic device (HoloLens 2) to execute the mixed reality–based collaborative system.

Results: A thorough experiment was conducted to validate the novel framework for mixed reality–based control of a collaborative robot. This framework was successfully applied to implement a collaborative system using a 5–degree of freedom robot (xArm-5) in a mixed reality environment. The framework was stable and worked smoothly throughout the collaborative session. Due to the distributed nature of cloud applications, there is a negligible latency between giving a command and the execution of the physical collaborative robot.

Conclusions: Opportunities for collaborative robots in telerehabilitation and teleoperation are vital as in any other field. The proposed framework was successfully applied in a collaborative session, and it can also be applied in other similar potential applications for robust and more promising performance.

(JMIR Biomed Eng 2022;7(1):e36734) doi:10.2196/36734

KEYWORDS

robot framework; mixed reality; collaborative robot; Unity; Windows Mixed Reality; Azure mixed reality services; HoloLens 2

Introduction

Background

Robots are becoming friends of humans by impacting and contributing to our daily life in many ways. They can carry out complex or repetitive activities for us, from household to industrial work, medicine, security, and so on [1]. Not only can robots assist in heavy works, but they can also help in education, especially during the COVID-19 pandemic. With the help of artificial intelligence, robots now can act as friends by monitoring behaviors and understanding our likes and dislikes.

They are offering various services effectively being used in different industries. Robots can boost the productivity of an industry by performing accurate, precise, fast, consistent, and high-quality work [2]. They can ensure safety by overtaking dangerous tasks in hazardous environments.

Collaborative robots, also known as cobots, are making the situation easier and more productive as they are designed to work in the same workspace as humans [3]. Both the robots and humans work delicately together, and robots assist human coworkers in completing different tasks [4]. This collaboration between humans and robots is changing the industrial production strategy. More industries are shifting to this manufacturing style daily because of flexibility, productivity, safety, reduced risk of injuries, and quality of performance in production [5]. Because of the recent advancement in the collaborative robots' application in various fields, the size of the global market in this field is growing daily. In 2018, the market value of this industry was around USD 649.1 million and was anticipated to expand to 44.5% by 2025 [6]. Therefore, the potential of collaborative robots, without any doubt, is huge, and more and more research is needed for the refinement of different approaches and applications in this field.

To support the needs of collaborative robots in different potential fields, researchers are working to propose different frameworks for different collaborative robot-based applications. Some researchers followed the agent-based system [7], whereas others also explored compliant control [8-10] and ergonomic aspects–based approach [11]. Still, researchers should consider other state-of-the-art technology-based approaches to designing a strong foundation for the huge potential of collaborative robots. One of the cutting-edge technologies in this era is mixed reality (MR), which blends both the digital and physical worlds to offer a promising solution for various applications [12]. MR merges virtual and augmented reality together, letting us incorporate the real world with digital data. Almost 150 firms in different fields have already adopted MR-based solutions, and it is estimated that by 2025 more than 14.4 million US employees will use smart glasses [13]. The possibility of mixed reality is huge, especially where human interaction is required. Therefore, the application will be promising if the mixed reality can successfully be applied in designing a framework for the collaborative robot.

In this study, a novel framework is proposed for the collaborative robot using mixed reality. Unity, Unity Hub, Windows Mixed Reality (WMR) platform, and Azure mixed reality services are adopted to design the framework and use a holographic device (HoloLens 2) to execute it. The framework can be used in various collaborative applications such as telerehabilitation or teleoperation.

The rest of the manuscript is outlined as follows: first, some recent advancements in this research area are briefly discussed; then, the development of the framework, along with the system architecture and the control of a collaborative robot with mixed reality, is represented; subsequently, one experiment using the framework and result of the experiment is illustrated; and finally, the conclusion of the study is portrayed.

Related Work

The collaborative robot is one of the potential fields in robotics, and researchers are working on the control system nowadays. Few researchers followed a unified approach by merging an impedance model with a dynamic model with friction and optimizing the assembly [14]. They used the proportional derivative control for the inner loop and impedance control for the outer loop. To evaluate the system's efficiency, a 6-DOF (Degrees of Freedom) series collaborative robot was used to perform peg-in-hole assembly tasks, and the performance of the system was accurate and flexible. One study proposed that by assessing the mobility of a collaborative robot, the performance of the robot system can be estimated [15]. Most of the available solutions are for a single robot, yet the model was evaluated using 3 automobiles. The model performed more competently than most of the strategies available. Another study represented the humans and robots as coworkers using geometric primitives, attraction vectors, and hypothetical repulsion by computing the distance and relative motion among them [16]. By applying this idea along with the robot's kinematic representation, the system achieved collision avoidance control by generating a nominal path to cautiously avoid collision with the human while performing the industrial operations. Repulsion-vector-reshaping was also introduced to ensure motion persistence, and the robot performed smoothly and successfully by avoiding collisions.

Rehabilitation robots and assistive robots are two potential applications for collaborative robots. Researchers are working on different approaches using skin surface electromyogram–based signals [17], nonlinear sliding mode control [18], geometric solution [19], and variable transformation for flatness geometric property [20] using collaborative robots to design robots for rehabilitation. Researchers have also followed learning latent actions from task demonstrations [21], reinforcement learning [22], digital image processing [23], and eye tracking–based assistive robot control [24] approaches for collaborative robots, focusing on assistive applications.

One study suggested controlling the momentum of a robot by considering the maximum speed acceptable to secure the safety of human coworkers [25]. The system estimated the allowable top velocity by using a collision model to predict the maximum force during a collision. The system enhanced the functionality and productivity of the collaborative tasks without risking human safety. Few researchers presented a new robotic system for collaborative robots blending mobile manipulators and supernumerary limbs [26]. Their robot could operate autonomously and be connected with humans as additional body parts. Other researchers presented a collaborative system that consists of hardware, software, and operational architecture of a humanoid robot trained with cognitive abilities [27]. The robot could identify the help a human coworker might need, recognize their activities, grasp objects, navigate, and so on. The experimental evaluation demonstrated that the robot performed safely and robustly while conducting collaborative tasks.

Researchers are also exploring framework-based approaches to construct generalized systems for different applications. In one

study, researchers proposed an open-source framework for a humanoid robot using cross architecture, which was low-cost in computation [28]. The framework was validated via both simulator and telemetry interface, and the result showed that it could be used to design new algorithms. In another study, researchers presented a framework for a collaborative human-robot environment using a commercial manipulator and their unique control method [10]. The framework included a trajectory planning and safety strategy by exploiting the human worker's experiences and was evaluated in a factory. In a similar study, a few researchers presented a framework for robot-assisted control in human-robot cooperation for a 7-DoF surgical robot [29]. The framework used manual motion to drive the tooltip, a 3D camera–based method to adjust the workspace, calculation of optimal instrument orientation, and cartesian interpolation to assure safety. In other studies, researchers proposed frameworks for different human-robot collaboration–based applications such as industrial cyber-physical systems [30], interaction in games (ie, Rock-Paper-Scissors) [31], and cooperative assembly duties [32].

The mixed reality–based approaches are becoming popular for different applications among researchers day by day as it has potential in many ways. Researchers used mixed reality–based methods to design an interface for human-robot interaction to teleoperate a robot [33] and a user interface to control teleoperated robotic platforms [34]. Researchers also used these approaches to design various robotic control systems. In one study, researchers developed an interface for human-robot communication using mixed reality for interactive robot manipulation control for mobile platforms [35]. The interface offered tools for robot path planning and visualized it for workers to comprehend robot behavior to ensure safety. The interface was successfully implemented and evaluated on Microsoft HoloLens. In another similar study, researchers used both mixed and virtual reality to design workspace for collaborative robots in the industry [36]. Robotic Operating System and Unity were used to design the system and tested in diverse settings. In another work, researchers presented an interactive control framework for both single and multirobot systems using mixed reality for various applications [37]. The system allowed interaction with robots by focusing on the visualization of their objective, and it could relate to any robots and mixed reality interfaces. The presented framework was evaluated experimentally, and the results verified the framework's capabilities in the mixed reality system.

Methods

Development of a Mixed Reality Framework for Robot Control

To develop the mixed reality–based system from scratch for controlling an assistive robot, some prerequisites and a few prior pieces of knowledge are required. Windows operating system–based computer and windows SDK (software development kit) with visual studio are needed to design the structure. The simplest approach to making mixed reality apps is installing either the Unity or Unreal game engines [38]. However, the same programs may be created for a custom engine using DirectX (Microsoft Corp). DirectX is a high-level interface that allows to directly access low-level functionality. It connects to Windows' hardware abstraction layer [39]. Unity is one of the most popular real time programming development platforms on the market, with C++-based runtime code and C#-based development scripting [40]. Unity and Unity Hub are used as cross-platform gaming engines and project management tools. The Mixed Reality Toolkit, if Unity is used, may be used for input simulation to test various input interactions, including hand-tracking and eye-tracking input. WMR is a Microsoft platform that debuted with Windows 10. The WMR platform enables developers to create apps that show digital material on holographic and virtual reality displays [41]. The Mixed Reality Feature Tool is needed to configure Unity while developing the framework. Interfacing, generating scenes, importing packages, and adding game objects to a scene all require a basic understanding of Unity. As the Unity scripts are written in C#, some fundamental C# expertise is also required. Any previous knowledge of NoSQL database systems and serverless functionalities will help with the system design. Finally, to implement the application, a holographic device (HoloLens 2) is required. By collecting and revealing digital information within the work setting and surroundings, Azure mixed reality services let people create, learn, and collaborate more efficiently [42]. Azure also helps secure and protect the stored data using 256-bit Advanced Encryption Standard and Transport Layer Security when data are in transit [43]. Azure mixed reality services bring 3D to mobile phones, headphones, and other devices that are not connected to the internet. Azure Remote Rendering and Azure Spatial Anchors are two mixed reality cloud technologies that let developers create captivating, immersive experiences across several platforms. These services enable incorporating spatial awareness into projects when developing 3D training, predictive equipment maintenance, and design review applications, all within the context of users' environments. The kinematics and dynamics of the assistive robot must be taken into consideration when developing the framework, which is explained below.

Assistive Robot

The xArm-5 robot is an end effector robot with 5 DoF, developed by UFactory [44]. To attain a high payload capability (3 kg) and appropriate repeatability of 0.1 mm, this robot is equipped with high-performance harmonic reducers and brushless motors. xArm-5 has a total workspace area of 0.7 meters. The Modbus Remote Terminal Unit protocol is used by xArm-5, and RS-485 communication is required to interact with a position, speed, or force control. These characteristics combine to make the xArm-5 one of the most versatile, high-precision, and multifunction robotic arms on the market. However, xArm-5 is comparatively heavyweight (11.2 kg) due to the big size of the motors; and due to its design features, this robot cannot be folded back during idle time (kinematic constraints). The graphical user interface for xArm-5 is provided by UFactory as xArm Studio and SDK for the languages Python, Robotic Operating System, and C++. Python SDK is used to control the xArm-5 for advanced capabilities in this research.

xArm-5 Robot's Kinematics and Dynamics

During the kinematic analysis, the modified Denavit-Hartenberg (DH) parameters are adopted to specify the xArm configuration of links and joints [45]. On the other hand, the iterative Newton-Euler method was applied for dynamic estimation to assess the joint torques corresponding to each activity of daily living.

Forward Kinematics

The link frame allocation (according to modified DH convention) of the xArm-5 robot is shown in Figure 1, where the yellow dots indicate that the direction is pointing into the viewing surface, cyan dot represents the direction pointing out of the viewing surface, and axis z defines the axis of rotation of each joint. To calculate the forward kinematics, the modified DH parameters corresponding to the link-frame allocation are given in Table 1. Moreover, Table 2 outlines the robot's link parameters.

Figure 1. Coordinate frame placement using the modified Denavit-Hartenberg parameters.

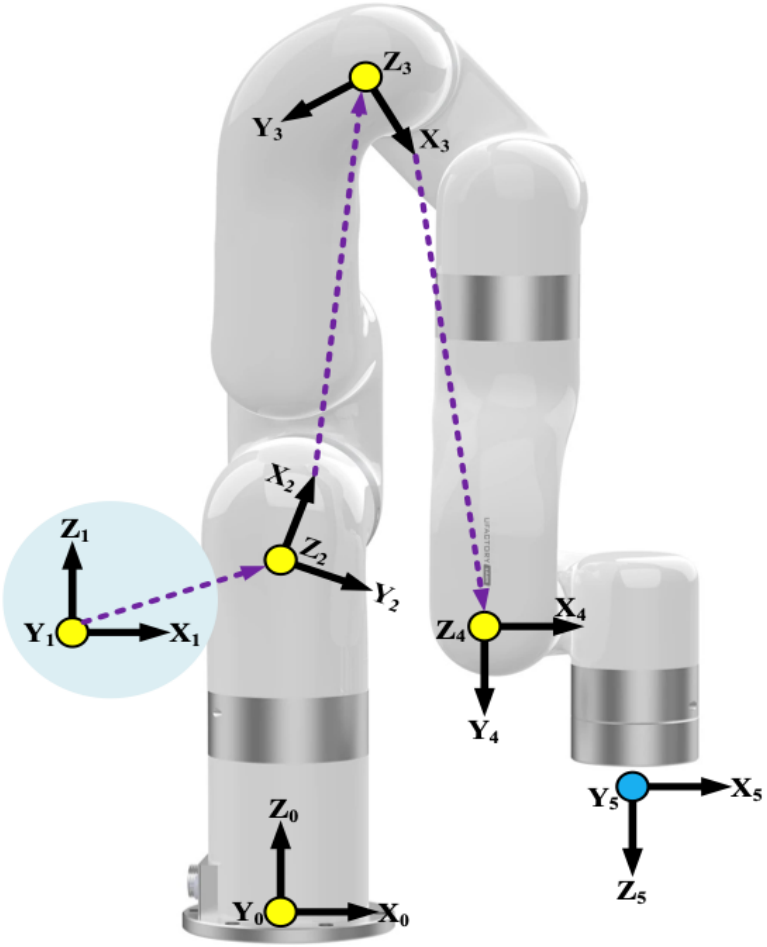


Table 1. Modified Denavit-Hartenberg parameters of xArm-5 robot.

1	0	–π/2	267	0
2	289.4886	0	0	–1.3849
3	351.1587	0	0	2.7331
4	76	–π/2	0	–1.3482
5	0	0	97	0

Here, α_i is the length of the common normal, α_i is the angle about common normal, d_i is the offset along the previous z-axis, and represents the joint angles. Note that the term represents the length of the link, and represents the offset of the angle. The values of those variables are shown in Table 2.

Table 2. Dimensional parameters of xArm-5 robots.

267 mm	284.5 mm	77.5 mm	342.5 mm	76 mm	–1.3849 rad	2.7331 rad	–1.3482 rad

The general form Homogeneous Transformation Matrix that relates two successive coordinate frames is represented by Equation (1).

Where  is the rotation matrix that represents the frame  in relation to frame  , and  is the vector that indicates the location of the origin of the frame  with respect to the frame  .

Moreover,  is the link twist,  corresponds to link length, d_i stands for link offset, and θ_i is the joint angle (radian) of the xArm5 Robot. The individual homogeneous transfer matrices that relate two successive frames of the xArm robot (Figure 1) are derived using Equation 1 and are given in Multimedia Appendix 1. The homogenous transformation matrix that relates frame {5} to frame {0} can be obtained by multiplying individual transformation matrices as expressed in Equation (2).

The single transformation matrix found from Equation (2) represents the reference frame's positions and orientations attached to the end effector with respect to the base reference frame {0}.

Dynamics of the xArm-5 Robot

The dynamic equation of the xARm5 Robot derived from the Newton-Euler formulation can be written in the following form:

Where  is the 5×5 mass matrix of the manipulator,  are the 5×1 acceleration vector,  is the 5×1 vector of centrifugal and Coriolis terms, and $G(\theta)$ is the 5×1 vector of gravity terms. Table 3 summarizes the mass/inertia parameters of the xArm-5 robot, and joint torques for the xArm-5 were calculated using Equation 3. Moreover, Table 4 presents the range of motion of each joint.

Table 3. Inertial parameters for each link of xArm-5 robot.

	Mass (kg)	Center of mass (mm)
Link 1	2.177	[0.15, 27.24, –13.57]
Link 2	2.011	[36.7, –220.9, 33.56]
Link 3	2.01	[68.34, 223.66, 1.1]
Link 4	1.206	[63.87, 29.3, 3.5]
Link 5	0.17	[0, –6.77, –10.98]

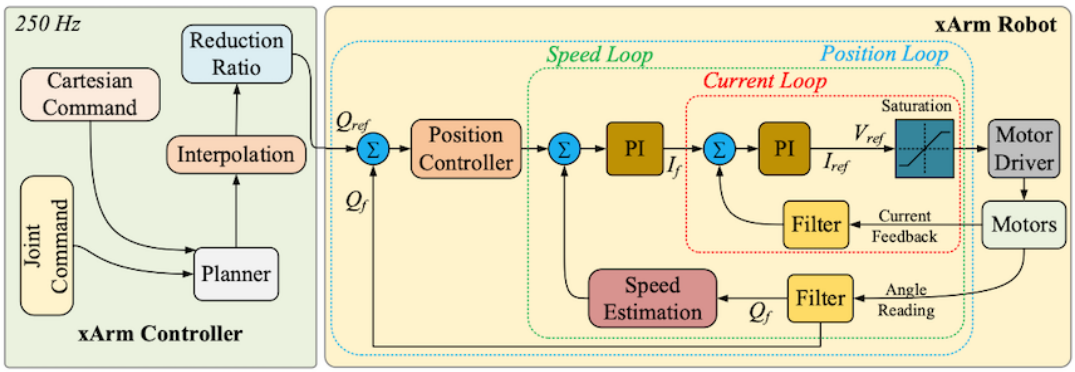
Table 4. Range of motion.

Joint	Working range
Joint number, deg	
Joint 1	±360
Joint 2	–118~120
Joint 3	–225~11
Joint 4	±360
Joint 5	–97~180
Maximum speed, deg/s	180

xArm-5's Control Architecture

Figure 2 illustrates the control architecture for the xArm-5 robot. The xArm controller generates two commands, the joints' torque and the cartesian, and updates the torque commands every 4 ms to execute in the xArm-5 controller. The torque commands are transformed to motor currents and finally to reference voltage for the motor drivers. The proportional integral controller is employed to acknowledge the real time control of the system. It also guarantees that the suitable control torque commands are transmitted to the joints and the reference voltage commands for the drivers. It also minimizes the differences between the desired and the measured currents.

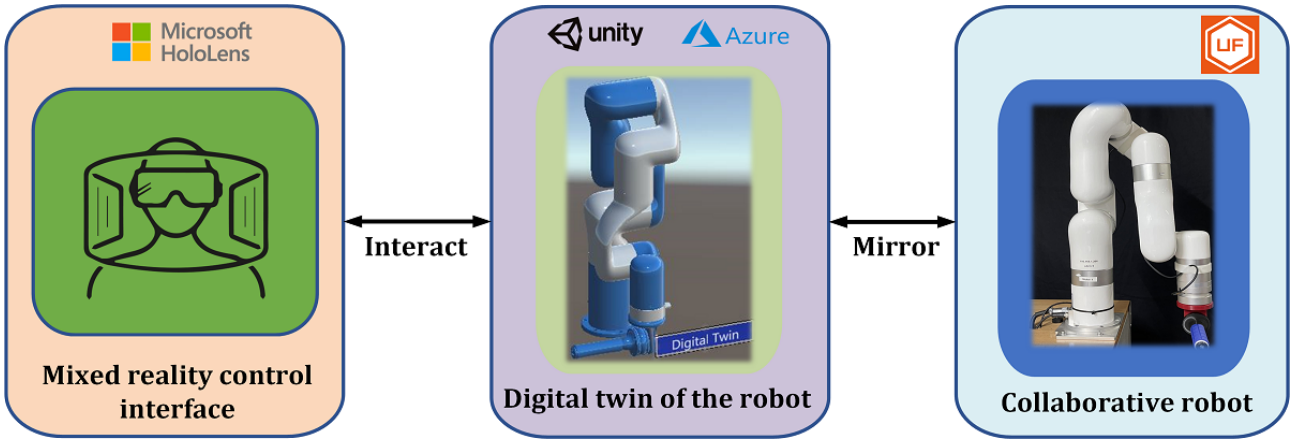
Figure 2. Control architecture of the system. I_f: filtered current; I_{ref}: reference current; PI: proportional integral; Q_f: filtered joint angle; Q_{ref}: reference joint angle; V_{ref}: reference voltage.



System Architecture

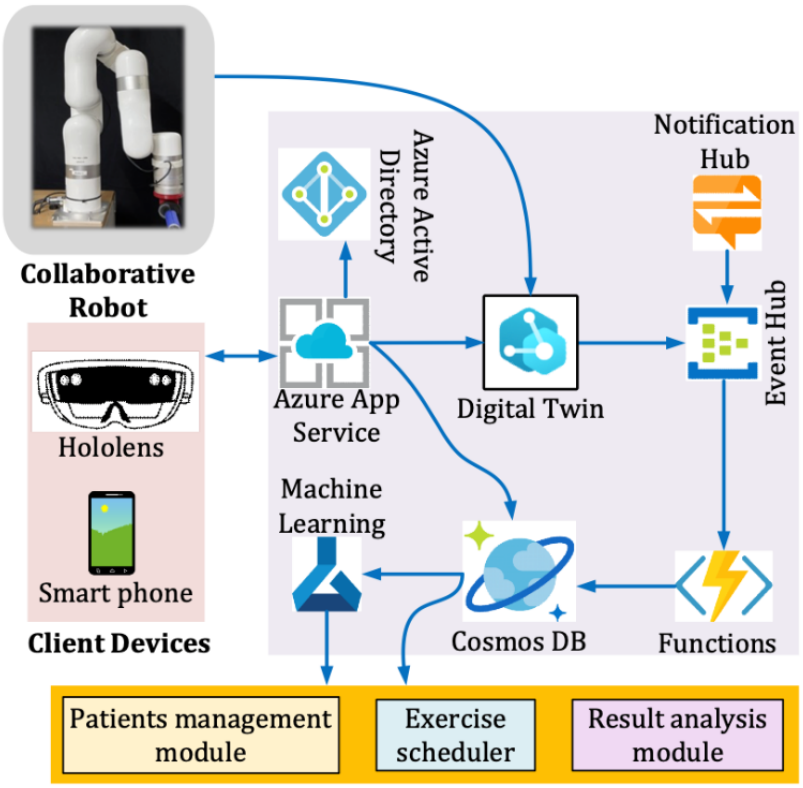
This section discusses the life cycle of a collaborative session involving a collaborative robot in a mixed reality environment. The user launches the application from their Mixed Reality headset, which is the application's access point (HoloLens 2). With the suitable digital twin and user, the provider can construct a collaborative session room. The user, as well as any allowed onlookers, can now join the collaboration session. Even though they are in separate places, all users in the collaboration session, including the host user (provider), are now in a digital collaboration session and communicating with the identical digital twin. Depending on the different needs, the provider can regulate the digital twin in several ways. Figure 3 shows the data flow of the connected system.

Figure 3. Data flow of the connected system.



Due to the distributed nature of cloud applications, the system can send a command to the digital twin, which the collaborative robot physically executes. All commands require a controller that runs on the robot's local network. WebSocket is a computer communications protocol that allows full-duplex communication channels over a single Transmission Control Protocol connection, which manages all device connections and interactions [46]. Users of the mixed reality collaborative session can now use high-level commands to control the collaborative robot. If the provider wants to control the robot to the desired path, it will be confined to moving along a fixed course and at varying distances along that trajectory. The user can provide input to the provider during this procedure, and key data such as the joint parameters and torques are collected and sent to the cloud. For example, using the Mixed Reality headset, the user can play interactive games that target specific muscles or motions to create a sense of confidence and accomplishment in their physical development. Thus, the entire quality of the therapy is improved. Internet of Things data can be stored and retrieved using the Azure cloud platform. Every collaborative robot sends telemetry data to a cloud platform, where it is stored indefinitely. In the case of teletherapy, relevant parameters such as patient range of motion and resistance to motion during various exercises are pushed to the cloud during each rehabilitation session. These parameters are used by the Azure cloud platform to support machine learning approaches that can adaptively develop appropriate rehabilitation strategies for each user. Figure 4 shows the details of the system architecture.

Figure 4. Detailed system architecture.



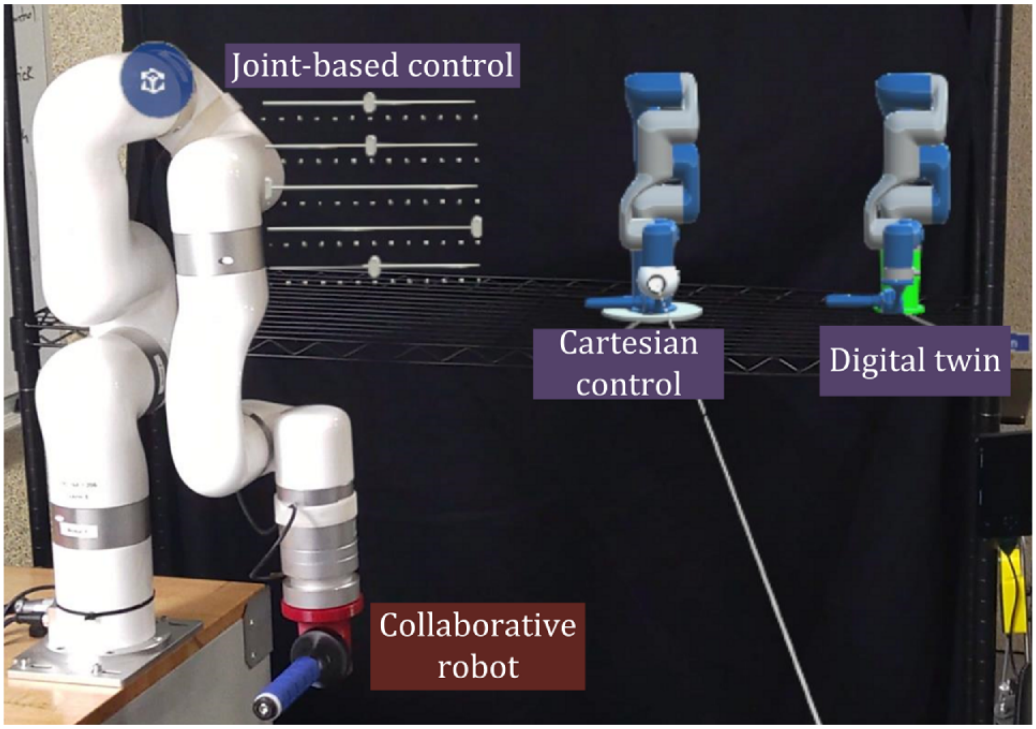
At first, in the event sequence of the mixed reality environment, the user launches the app on the client device and connects with Azure App Service via WebSocket. Then, the user is authenticated via Azure Active Directory. Subsequently, the user can select a digital twin model they wish to interact with. The App Service will retrieve assets corresponding to the selected digital twin, including 3D models and data. Next, the App Service provides a user with their requested data. The Digital Twin pushes incoming data to the Event Hub when the system is running, which fires a serverless function on the Azure cloud server. The serverless function updates database values. At the final step of data flow, some machine learning models will be deployed, which will use historical users' data and real time collaboration metrics for future automated result analysis.

Control of a Collaborative Robot Through the Mixed Reality Environment

The Azure cloud platform enables multiple users to join a shared collaboration session. In this collaboration session, users can visualize and interact with one digital twin. The digital twin will move in tandem with the physical robot and mirror its movements. Users in the collaboration session can additionally send high-level commands to control the digital twin.

Figure 5 shows the mixed reality user interface of the proposed framework. During the collaboration session, a user can control a collaborative robot in several ways, such as joint angle control, cartesian control, and preplanned trajectory control. The virtual sliders are used in joint-based control. On the other hand, in cartesian mode end effector is controlled by moving the virtual end effector in a mixed reality environment. Furthermore, the provider can set a desired path to follow for a collaborative robot in the preplanned trajectory control mode.

Figure 5. Mixed reality interface.



Ethical Considerations

The researchers involved in the project took part in the experiments to demonstrate the proof of concept of the teleoperation using the proposed framework. Therefore, ethical approval is not required for this study. The project did not focus on any intervention development or intervention efficacy testing; hence, we did not recruit any participants and did not seek ethics approval for this project.

Results

To validate the proposed framework, an end effector type 5 DoF assistive robot is mounted on a table (Figure 6a). The purpose of this assistive robot is to give therapy with a pretrained trajectory and by a practitioner. The designed app, which contains the mixed reality interface, is deployed to the HoloLens 2 to control the assistive robot. A practitioner can wear the HoloLens 2 and control the therapeutic sessions remotely in a mixed reality environment (Figure 6b). Figure 6c depicts a collaboration session between a practitioner and a patient where the robot can be controlled and monitored remotely in a mixed reality environment using the proposed framework.

The system provides an overall cross-platform rehabilitation experience for its users. A clinician can remotely assist a patient in both active and passive rehabilitation therapy via a shared mixed reality–based collaboration session. The rehabilitation robot that exists locally for the patient has a digital twin that lives on the Azure cloud platform. A therapist or clinician can interact with this virtual digital twin and use it to assist with rehabilitation therapy. Figure 7 shows a mixed reality interface that visualizes robot data such as torque and temperature for each joint, as well as force sensor data. In this manner, the therapist can visualize the key metrics that estimate the patients' improvement, such as range of motion and resistance. They can use these data to recommend optimal rehabilitation plans for a patient. Users in the collaboration session can summon data visualizations of both real time and historical data.

Figure 6. Experimental setup for the proposed collaborative robot framework: (a) xArm-5 Robot as the collaborative robot; (b) user with HoloLens 2 headset; and (c) collaborative session via the mixed reality–based framework (Multimedia Appendix 2).

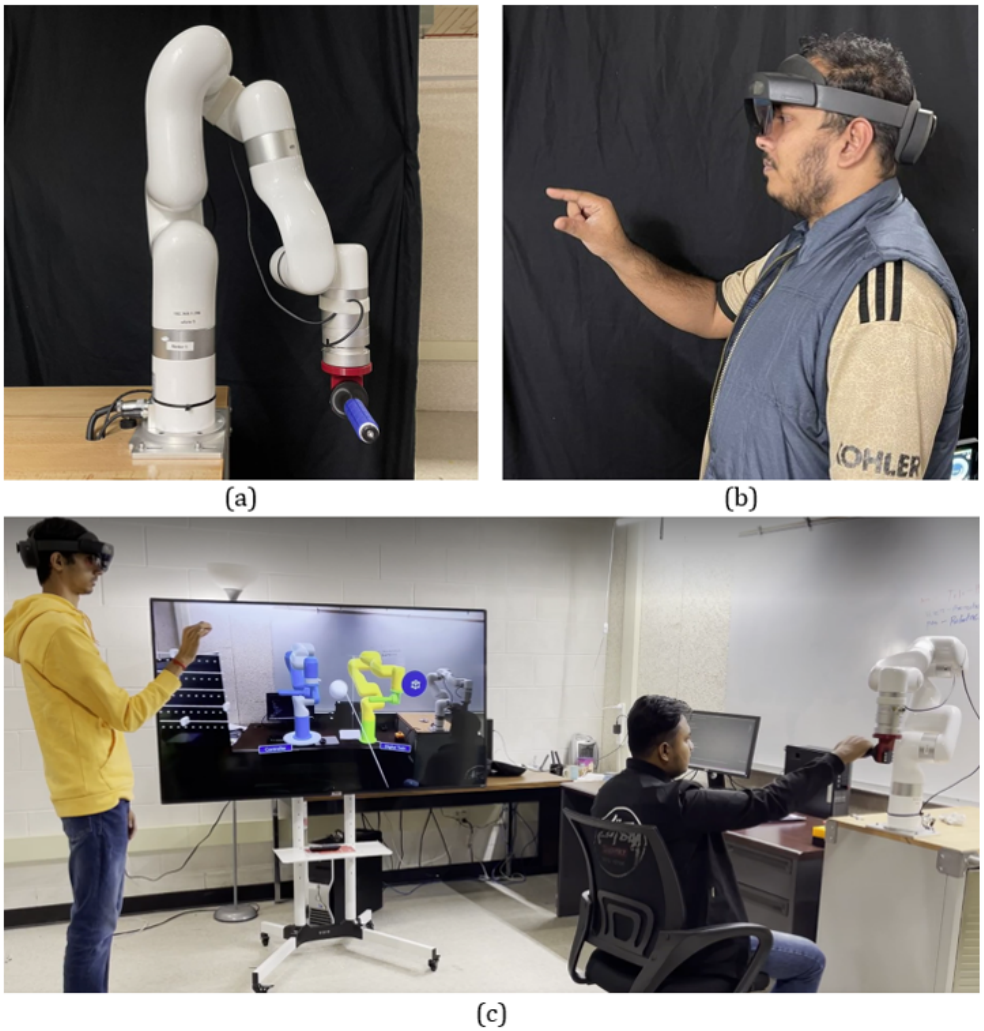
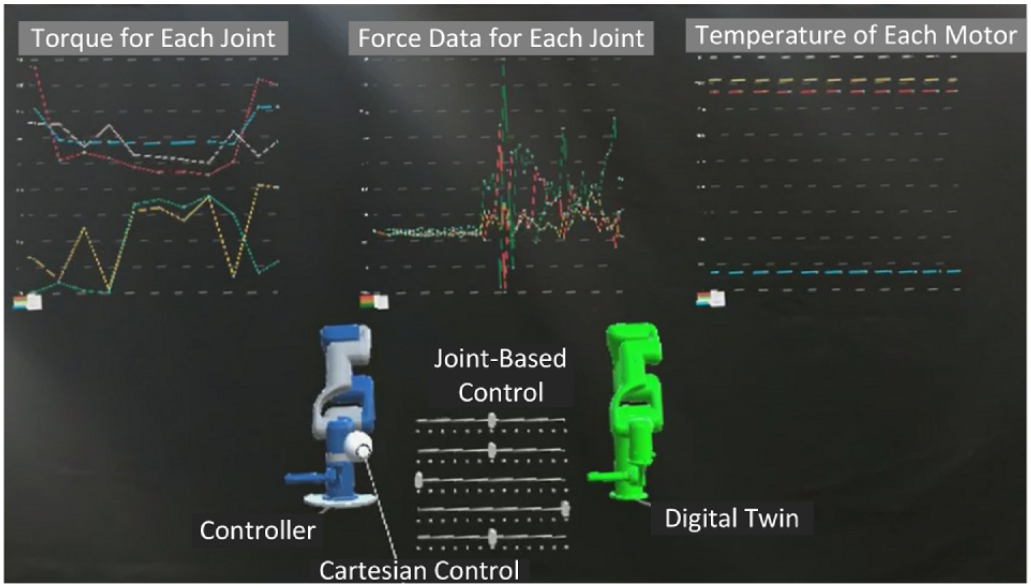


Figure 7. Visualizing the key metrics in the interactive mixed reality environment.



Discussion

Framework for collaborative robots has many potential applications in industry and telehealth such as teleoperation and telerehabilitation. Especially, a pandemic situation such as COVID-19 has affected all aspects of health care and gave the realization of the need for telehealth, which can help health care workers and patients to protect themselves from the risk of disease transmission. Telehealth can also offer personalized rehabilitation programs, real time control, continuous interaction with doctors, negligible waiting time, and so on. Using the proposed framework, it will be handy to implement different systems for teleoperation. While the proposed mixed reality–based framework promises a stable control system for collaborative robots, there are some limitations to it [47-49]. To use the framework, there should be a continuous and stable connection. The system becomes unstable and inoperable if the connection is lost. Security of personal health data is also a concern. Furthermore, the system is expensive compared to other regular rehabilitation, as a holographic device and a collaborative robot is needed for this [50]. It is worth mentioning that a holographic device such as HoloLens 2 should not be worn for extended periods of time. Possible side effects of HoloLens include headache, dizziness, or loss of balance [50,51]. It is important to use these tools responsibly to achieve the maximum benefit from the services they provide. To improve the framework, in the future, the focus will be given to stable communication, personalized rehabilitation program, and real time control and monitoring by the expert.

Conclusion

Collaborative robots and their applications have a magnificent impact in the rehabilitation and teleoperation fields. A framework for collaborative robots is very much needed to meet the demands of collaborative robots. The proposed mixed reality–based framework for collaborative robots can be used for different telehealth applications such as teleoperation and telerehabilitation and can guide other researchers to conduct further studies for the advancement of humans. Several state-of-the-art technologies were used while developing the framework, including Unity, WMR platform, Azure mixed reality services, and HoloLens 2. The framework is validated by conducting a comprehensive, collaborative experiment using a 5-DoF collaborative robot (xArm-5) in a mixed reality environment. In the future, the study will continue with mixed reality–based applications for collaborative robots. Further research will be conducted on telerehabilitation and teleoperation to design a more robust and stable framework for further advancement.

Acknowledgments

This research is partially funded by the Connected System Institute, University of Wisconsin-Milwaukee (Grant AAH9481).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Individual homogeneous transfer matrix.
[DOCX File , 14 KB - biomedeng_v7i1e36734_app1.docx]

Multimedia Appendix 2
Experiment of collaborative session via mixed reality–based framework.
[MOV File , 73149 KB - biomedeng_v7i1e36734_app2.mov]

References

1. Muthoni J. Use of Robotics in Our Daily Lives. Jonas Muthoni. 2021 Jun 16. URL: https://jonasmuthoni.com/blog/use-of-robotics-in-daily-lives/ [accessed 2022-01-02]
2. Benefits of Robots. RobotWorx. 2021. URL: https://www.robots.com/articles/benefits-of-robots [accessed 2022-01-02]
3. Why Cobots? Universal Robots. 2021. URL: https://www.universal-robots.com/products/collaborative-robots-cobots-benefits [accessed 2022-01-02]
4. Khasis D. How Humans and Robots Can Work Together for Better Warehouse Management. RIS News. 2019. URL: https://risnews.com/how-humans-and-robots-can-work-together-better-warehouse-management [accessed 2022-01-02]
5. Human-Robot Collaboration: 3 Case Studies. Wevolver. 2020. URL: https://www.wevolver.com/article/humanrobot.collaboration.3.case.studies [accessed 2022-01-02]
6. Grand View Research 2019. URL: https://www.grandviewresearch.com/industry-analysis/collaborative-robots-market [accessed 2022-01-02]
7. Schou C, Madsen O. A plug and produce framework for industrial collaborative robots. International Journal of Advanced Robotic Systems 2017 Jul 17;14(4):172988141771747. [doi: 10.1177/1729881417717472]

8. Brahmi B, Laraki MH, Saad M, Rahman M, Ochoa-Luna C, Brahmi A. Compliant adaptive control of human upper-limb exoskeleton robot with unknown dynamics based on a Modified Function Approximation Technique (MFAT). Robotics and Autonomous Systems 2019 Jul;117:92-102. [doi: 10.1016/j.robot.2019.02.017]
9. Brahmi B, Saad M, Brahmi A, Luna CO, Rahman MH. Compliant control for wearable exoskeleton robot based on human inverse kinematics. International Journal of Advanced Robotic Systems 2018 Nov 22;15(6):172988141881213. [doi: 10.1177/1729881418812133]
10. Maric B, Mutka A, Orsag M. Collaborative Human-Robot Framework for Delicate Sanding of Complex Shape Surfaces. IEEE Robot. Autom. Lett 2020 Apr;5(2):2848-2855. [doi: 10.1109/lra.2020.2969951]
11. Kim W, Peternel L, Lorenzini M, Babič J, Ajoudani A. A Human-Robot Collaboration Framework for Improving Ergonomics During Dexterous Operation of Power Tools. Robotics and Computer-Integrated Manufacturing 2021 Apr;68:102084. [doi: 10.1016/j.rcim.2020.102084]
12. What is mixed reality? Microsoft. 2022 Apr 28. URL: https://docs.microsoft.com/en-us/windows/mixed-reality/discover/mixed-reality [accessed 2022-01-02]
13. Importance of Mixed Reality in Real World. AREA. 2020. URL: https://thearea.org/ar-news/importance-of-mixed-reality-in-real-world/ [accessed 2022-01-02]
14. Zeng F, Xiao J, Liu H. Force/Torque Sensorless Compliant Control Strategy for Assembly Tasks Using a 6-DOF Collaborative Robot. IEEE Access 2019;7:108795-108805. [doi: 10.1109/access.2019.2931515]
15. Grigore LS, Priescu I, Joita D, Oncioiu I. The Integration of Collaborative Robot Systems and Their Environmental Impacts. Processes 2020 Apr 23;8(4):494. [doi: 10.3390/pr8040494]
16. Safeea M, Neto P, Bearee R. On-line collision avoidance for collaborative robot manipulators by adjusting off-line generated paths: An industrial use case. Robotics and Autonomous Systems 2019 Sep;119:278-288. [doi: 10.1016/j.robot.2019.07.013]
17. Kiguchi K, Rahman MH, Sasaki M, Teramoto K. Development of a 3DOF mobile exoskeleton robot for human upper-limb motion assist. Robotics and Autonomous Systems 2008 Aug;56(8):678-691. [doi: 10.1016/j.robot.2007.11.007]
18. Rahman MH, Kittel-Ouimet T, Saad M, Kenné J, Archambault PS. Development and Control of a Robotic Exoskeleton for Shoulder, Elbow and Forearm Movement Assistance. Applied Bionics and Biomechanics 2012;9(3):275-292. [doi: 10.1155/2012/956310]
19. Assad-Uz-Zaman M, Islam M, Rahman M, Wang Y, McGonigle E. Kinect Controlled NAO Robot for Telerehabilitation. Journal of Intelligent Systems 2020 2020 Jul 28;30(1):224-239. [doi: 10.1515/jisys-2019-0126]
20. Brahmi B, Ahmed T, Bojairami IE, Swapnil AAZ, Assad-Uz-Zaman M, Schultz K, et al. Flatness Based Control of a Novel Smart Exoskeleton Robot. IEEE/ASME Trans. Mechatron 2022 Apr;27(2):974-984. [doi: 10.1109/tmech.2021.3076956]
21. Losey DP, Jeon HJ, Li M, Srinivasan K, Mandlekar A, Garg A, et al. Learning latent actions to control assistive robots. Auton Robots 2022 Aug 04;46(1):115-147 [FREE Full text] [doi: 10.1007/s10514-021-10005-w] [Medline: 34366568]
22. Erickson Z, Gangaram V, Kapusta A, Liu C, Kemp C. Assistive Gym: A Physics Simulation Framework for Assistive Robotics. 2020 Presented at: IEEE International Conference on Robotics and Automation (ICRA); May 31-August 31, 2020; Paris, France. [doi: 10.1109/icra40945.2020.9197411]
23. Sunny MSH, De Caro JS, Rulik I, Zarif MII, Rahman M, Wang I, et al. Nose Tracking Assistive Robot Control for the People with Motor Dysfunctions. Archives of Physical Medicine and Rehabilitation 2021 Oct;102(10):e82-e83. [doi: 10.1016/j.apmr.2021.07.718]
24. Sunny MSH, Zarif MII, Rulik I, Sanjuan J, Rahman MH, Ahamed SI, et al. Eye-gaze control of a wheelchair mounted 6DOF assistive robot for activities of daily living. J Neuroeng Rehabil 2021 Dec 18;18(1):173 [FREE Full text] [doi: 10.1186/s12984-021-00969-2] [Medline: 34922590]
25. Shin H, Seo K, Rhim S. Allowable maximum safe velocity control based on human-robot distance for collaborative robot. 2018 Presented at: IEEE International Conference on Ubiquitous Robots (UR); 26-30 June 2018; Honolulu, HI, USA p. 401-405. [doi: 10.1109/urai.2018.8441887]
26. Kim W, Balatti P, Lamon E, Ajoudani A. MOCA-MAN: A MObile and reconfigurable Collaborative Robot Assistant for conjoined huMAN-robot actions. 2020 Presented at: IEEE International Conference on Robotics and Automation (ICRA); May 31-August 31, 2020; Paris, France. [doi: 10.1109/icra40945.2020.9197115]
27. Asfour T, Kaul L, Wächter M, Ottenhaus S, Weiner P, Rader S, et al. Armar-6: A collaborative humanoid robot for industrial environments. 2018 Presented at: IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids); November 6-9, 2018; Beijing, China p. 447-454. [doi: 10.1109/humanoids.2018.8624966]
28. Perico DH, Homem TPD, Almeida AC, Silva IJ, Vilão CO, Ferreira VN, et al. Humanoid Robot Framework for Research on Cognitive Robotics. J Control Autom Electr Syst 2018 May 31;29(4):470-479. [doi: 10.1007/s40313-018-0390-y]
29. Sandoval J, Su H, Vieyres P, Poisson G, Ferrigno G, De Momi E. Collaborative framework for robot-assisted minimally invasive surgery using a 7-DoF anthropomorphic robot. Robotics and Autonomous Systems 2018 Aug;106:95-106. [doi: 10.1016/j.robot.2018.04.001]
30. Khalid A, Kirisci P, Khan ZH, Ghrairi Z, Thoben K, Pannek J. Security framework for industrial collaborative robotic cyber-physical systems. Computers in Industry 2018 May;97:132-145. [doi: 10.1016/j.compind.2018.02.009]
31. Brock H, Ponce Chulani J, Merino L, Szapiro D, Gomez R. Developing a Lightweight Rock-Paper-Scissors Framework for Human-Robot Collaborative Gaming. IEEE Access 2020;8:202958-202968. [doi: 10.1109/access.2020.3033550]

32. Sadrfaridpour B, Wang Y. Collaborative Assembly in Hybrid Manufacturing Cells: An Integrated Framework for Human–Robot Interaction. IEEE Trans. Automat. Sci. Eng 2018 Jul;15(3):1178-1192. [doi: 10.1109/tase.2017.2748386]
33. Cousins M, Yang C, Chen J, He W, Ju Z. Development of a mixed reality based interface for human robot interaciotn. 2017 Presented at: IEEE International Conference on Machine Learning and Cybernetics (ICMLC) ;1; July 9-12, 2017; Ningbo, China p. 27-34. [doi: 10.1109/icmlc.2017.8107738]
34. Cancedda L, Cannavò A, Garofalo G, Lamberti F, Montuschi P, Paravati G. Mixed reality-based user interaction feedback for a hand-controlled interface targeted to robot teleoperation. 2017 Presented at: International Conference on Augmented Reality, Virtual Reality and Computer Graphics ;10325; June 12-15, 2017; Ugento, Italy p. 447-463. [doi: 10.1007/978-3-319-60928-7_38]
35. Ostanin M, Yagfarov R, Klimchik A. Interactive Robots Control Using Mixed Reality. IFAC-PapersOnLine 2019;52(13):695-700. [doi: 10.1016/j.ifacol.2019.11.307]
36. Siegele D, Steiner D, Giusti A, Riedl M, Matt D. Optimizing Collaborative Robotic Workspaces in Industry by Applying Mixed Reality. 2021 Presented at: International Conference on Augmented Reality, Virtual Reality and Computer Graphics 2021;12980. Springer, Cham; September 7-10, 2021; Virtual Event p. 544-559. [doi: 10.1007/978-3-030-87595-4_40]
37. Ostanin M, Yagfarov R, Devitt D, Akhmetzyanov A, Klimchik A. Multi robots interactive control using mixed reality. International Journal of Production Research 2020 Nov 07;59(23):7126-7138. [doi: 10.1080/00207543.2020.1834640]
38. Choosing your engine. Microsoft. URL: https://docs.microsoft.com/en-us/windows/mixed-reality/develop/choosing-an-engine?tabs=unity [accessed 2022-01-02]
39. Download DirectX End-User Runtime Web Installer from Official Microsoft Download Center. Microsoft. URL: https://www.microsoft.com/en-us/download/details.aspx?id=35 [accessed 2022-01-02]
40. Unity Real-Time Development Platform. Unity. URL: https://unity.com/ [accessed 2022-01-02]
41. Configure your project without MRTK. Microsoft. 2021. URL: https://docs.microsoft.com/en-us/windows/mixed-reality/develop/unity/configure-unity-project [accessed 2022-01-02]
42. Azure mixed reality cloud services overview. Microsoft. URL: https://docs.microsoft.com/en-us/windows/mixed-reality/develop/mixed-reality-cloud-services [accessed 2022-01-02]
43. Data Privacy in the Trusted Cloud. Microsoft Azure. URL: https://azure.microsoft.com/en-us/overview/trusted-cloud/privacy/ [accessed 2022-01-12]
44. UFACTORY xArm 5 Lite. UFACTORY. URL: https://www.ufactory.cc/products/xarm-5-lite-2020 [accessed 2022-01-02]
45. Kovalchuk A. Modified Denavit-Hartenberg Coordinate System for Robot Actuating Mechanisms with Tree-like Kinematic Structure. S&E BMSTU 2014 Dec 03;15(11):244. [doi: 10.7463/1115.0826673]
46. What is web socket and how it is different from the HTTP? GeeksforGeeks. 2022 Feb 21. URL: https://www.geeksforgeeks.org/what-is-web-socket-and-how-it-is-different-from-the-http/ [accessed 2022-01-02]
47. Van Krevelen D, Poelman R. A Survey of Augmented Reality Technologies, Applications and Limitations. IJVR 2010 Jan 01;9(2):1-20. [doi: 10.20870/ijvr.2010.9.2.2767]
48. Filipenko M, Angerer A, Reif W. Opportunities and limitations of mixed reality holograms in industrial robotics. arXiv preprint 2020 Jan 22. [doi: 10.48550/arXiv.2001.08166]
49. Takata T, Nakabayashi S, Kondo H, Yamamoto M, Furui S, Shiraishi K, et al. Mixed Reality Visualization of Radiation Dose for Health Professionals and Patients in Interventional Radiology. J Med Syst 2021 Feb 17;45(4):38 [FREE Full text] [doi: 10.1007/s10916-020-01700-9] [Medline: 33594609]
50. Morimoto T, Kobayashi T, Hirata H, Otani K, Sugimoto M, Tsukamoto M, et al. XR (Extended Reality: Virtual Reality, Augmented Reality, Mixed Reality) Technology in Spine Medicine: Status Quo and Quo Vadis. J Clin Med 2022 Jan 17;11(2):470 [FREE Full text] [doi: 10.3390/jcm11020470] [Medline: 35054164]
51. Condino S, Turini G, Parchi PD, Viglialoro RM, Piolanti N, Gesi M, et al. How to Build a Patient-Specific Hybrid Simulator for Orthopaedic Open Surgery: Benefits and Limits of Mixed-Reality Using the Microsoft HoloLens. J Healthc Eng 2018 Nov 01;2018:5435097-5435012 [FREE Full text] [doi: 10.1155/2018/5435097] [Medline: 30515284]

Abbreviations

DH: Denavit-Hartenberg
DOF: Degrees of Freedom
MR: Mixed Reality
SDK: software development kit
WMR: Windows Mixed Reality

Edited by G Eysenbach; submitted 23.01.22; peer-reviewed by MR Islam, N Silva; comments to author 30.03.22; revised version received 13.04.22; accepted 26.04.22; published 17.05.22.

Please cite as:
Shahria MT, Sunny MSH, Zarif MII, Khan MMR, Modi PP, Ahamed SI, Rahman MH
A Novel Framework for Mixed Reality–Based Control of Collaborative Robot: Development Study
JMIR Biomed Eng 2022;7(1):e36734
URL: https://biomedeng.jmir.org/2022/1/e36734
doi:10.2196/36734
PMID:

Original Paper

The Classification of Abnormal Hand Movement to Aid in Autism Detection: Machine Learning Study

Anish Lakkapragada[1]; Aaron Kline[1], BS; Onur Cezmi Mutlu[2], MS; Kelley Paskov[3], MS; Brianna Chrisman[4], MS; Nathaniel Stockham[5], MS; Peter Washington[6], PhD; Dennis Paul Wall[3], PhD

[1]Division of Systems Medicine, Department of Pediatrics, Stanford University, Stanford, CA, United States

[2]Department of Electrical Engineering, Stanford University, Stanford, CA, United States

[3]Department of Biomedical Data Science, Stanford University, Stanford, CA, United States

[4]Department of Bioengineering, Stanford University, Stanford, CA, United States

[5]Department of Neuroscience, Stanford University, Stanford, CA, United States

[6]Information and Computer Sciences, University of Hawai'i at Mānoa, Honolulu, HI, United States

Corresponding Author:
Peter Washington, PhD
Information and Computer Sciences
University of Hawai'i at Mānoa
2500 Campus Rd
Honolulu, HI, 96822
United States
Phone: 1 5126800926
Email: peter.y.washington@hawaii.edu

Abstract

Background: A formal autism diagnosis can be an inefficient and lengthy process. Families may wait several months or longer before receiving a diagnosis for their child despite evidence that earlier intervention leads to better treatment outcomes. Digital technologies that detect the presence of behaviors related to autism can scale access to pediatric diagnoses. A strong indicator of the presence of autism is self-stimulatory behaviors such as hand flapping.

Objective: This study aims to demonstrate the feasibility of deep learning technologies for the detection of hand flapping from unstructured home videos as a first step toward validation of whether statistical models coupled with digital technologies can be leveraged to aid in the automatic behavioral analysis of autism. To support the widespread sharing of such home videos, we explored privacy-preserving modifications to the input space via conversion of each video to hand landmark coordinates and measured the performance of corresponding time series classifiers.

Methods: We used the Self-Stimulatory Behavior Dataset (SSBD) that contains 75 videos of hand flapping, head banging, and spinning exhibited by children. From this data set, we extracted 100 hand flapping videos and 100 control videos, each between 2 to 5 seconds in duration. We evaluated five separate feature representations: four privacy-preserved subsets of hand landmarks detected by MediaPipe and one feature representation obtained from the output of the penultimate layer of a MobileNetV2 model fine-tuned on the SSBD. We fed these feature vectors into a long short-term memory network that predicted the presence of hand flapping in each video clip.

Results: The highest-performing model used MobileNetV2 to extract features and achieved a test F1 score of 84 (SD 3.7; precision 89.6, SD 4.3 and recall 80.4, SD 6) using 5-fold cross-validation for 100 random seeds on the SSBD data (500 total distinct folds). Of the models we trained on privacy-preserved data, the model trained with all hand landmarks reached an F1 score of 66.6 (SD 3.35). Another such model trained with a select 6 landmarks reached an F1 score of 68.3 (SD 3.6). A privacy-preserved model trained using a single landmark at the base of the hands and a model trained with the average of the locations of all the hand landmarks reached an F1 score of 64.9 (SD 6.5) and 64.2 (SD 6.8), respectively.

Conclusions: We created five lightweight neural networks that can detect hand flapping from unstructured videos. Training a long short-term memory network with convolutional feature vectors outperformed training with feature vectors of hand coordinates and used almost 900,000 fewer model parameters. This study provides the first step toward developing precise deep learning methods for activity detection of autism-related behaviors.

(JMIR Biomed Eng 2022;7(1):e33771) doi:10.2196/33771

KEYWORDS

deep learning; machine learning; activity recognition; applied machine learning; landmark detection; autism; diagnosis; health informatics; detection; feasibility; video; model; neural network

Introduction

Autism affects almost 1 in 44 people in America [1] and is the fastest growing developmental delay in the United States [2,3]. Although autism can be identified accurately by 24 months of age [4,5], the average age of diagnosis is slightly below 4.5 years [6]. This is problematic because earlier intervention leads to improved treatment outcomes [7]. Mobile digital diagnostics and therapeutics can help bridge this gap by providing scalable and accessible services to underserved populations lacking access to care. The use of digital and mobile therapies to support children with autism has been explored and validated in wearable devices [8-15] and smartphones [16-22] enhanced by machine learning models to help automate and streamline the therapeutic process.

Mobile diagnostic efforts for autism using machine learning have been explored in prior literature. Autism can be classified with high performance using 10 or fewer behavioral features [23-28]. While some untrained humans can reliably distinguish these behavioral features [25,29-36], an eventual goal is to move away from human-in-the-loop solutions toward automated and privacy-preserving diagnostic solutions [37,38]. Preliminary efforts in this space have included automated detection of autism-related behaviors such as head banging [39], emotion evocation [40-42], and eye gaze [43].

Restrictive and repetitive movement such as hand stimming is a primary behavioral feature used by diagnostic instruments for autism [44]. Because computer vision classifiers for abnormal hand movement do not currently exist, at least in the public domain, we strived to create a classifier that can detect this autism-related feature as a first step toward automated clinical support systems for developmental delays like autism.

Pose estimation and activity recognition have been explored as a method for detection of self-stimulatory behaviors. Vyas et al [45] retrained a 2D Mask region-based convolutional neural network (R-CNN) [46] to obtain the coordinates of 15 body landmarks that were then transformed into a Pose Motion (PoTion) representation [47] and fed to a convolutional neural network (CNN) model for a prediction of autism-related atypical movements. This approach resulted in a 72.4% classification accuracy with 72% precision and 92% recall. Rajagopalan and Goecke [48] used the Histogram of Dominant Motions (HDM) representation to train a model to detect self-stimulatory behaviors [48]. On the Self-Stimulatory Behavior Dataset (SSBD) [49], which we also used in this study, the authors achieved 86.6% binary accuracy when distinguishing head banging versus spinning and 76.3% accuracy on the 3-way task of distinguishing head banging, spinning, and hand flapping. We note that they did not train a classifier with a control class absent of any self-stimulatory behavior. Zhao et al [50] used head rotation range and rotations per minute in the yaw, pitch, and roll directions as features for autism detection classifiers. This reached 92.11% classification accuracy with a decision tree model that used the head rotation range in the roll direction and the amount of rotations per minute in the yaw direction as features.

Building upon these prior efforts, we developed a computer vision classifier for abnormal hand movement displayed by children. In contrast to prior approaches to movement-based detection of autism, which use extracted activity features to train a classifier to detect autism directly, we aim to detect autism-related behaviors that may contribute to an autism diagnosis but that may also be related to other behavioral symptoms. We trained our abnormal hand movement classifier on the SSBD, as it is the only publicly available data set of videos depicting abnormal hand movement in children. We used cross-validation and achieved an F1 score of 84% using convolutional features emitted per frame by a fine-tuned MobileNetV2 model fed into a long short-term memory (LSTM). We also explored privacy-preserving hand-engineered feature representations that may support the widespread sharing of home videos.

Methods

Overview

We compared five separate training approaches: four subsets of MediaPipe hand landmarks fed into an LSTM and fine-tuned MobileNetV2 convolutional features fed into an LSTM. The hand landmark approaches provided an exploration of activity detection on privacy-preserved feature representations. Because we strived to use machine learning classifiers in low-resource settings such as mobile devices, we additionally aimed to make our models and feature representations as light as possible.

Data Set

We used the SSBD [49] for training and testing of our models. To the best of our knowledge, SSBD is the only publicly available data set of self-stimulatory behaviors containing examples of head banging, hand flapping, and spinning. SSBD includes the URLs of 75 YouTube videos, and for each video, annotations of the time periods (eg, second 1 to second 35) when each self-stimulatory behavior was performed. Multiple videos contain multiple time periods for the same behavior (eg, seconds 1-3 and 5-9 both contain hand flapping) as well as multiple behaviors (eg, seconds 1-3 show head banging and seconds 5-9 show hand flapping). We only used the hand flapping annotations.

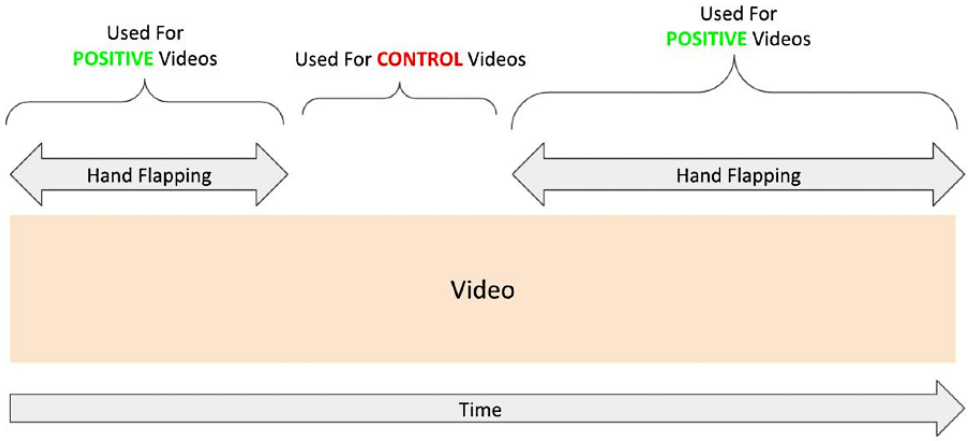
Preprocessing

To obtain control videos absent of hand flapping displays, we first downloaded all YouTube videos in SSBD that contained sections of hand flapping. Each section in a video exhibiting hand flapping was extracted to create a new clip. The parts of

the video without hand flapping (ie, with no annotations) were isolated to create control clips. This data curation process is illustrated in Figure 1.

After extracting all positive and control clips from the downloaded videos, we aimed to maximize the amount of training data in each class. Because a hand flapping event occurs within a couple of seconds, we split any clips longer than 2 seconds into smaller clips. We manually deleted any videos that were qualitatively shaky or of low quality. In total, we extracted 50 video clips displaying hand flapping and 50 control videos.

Figure 1. Extraction of positive and control videos. Sections of a video demonstrating hand flapping are separated to create positive videos, and segments between the hand flapping sections are used as control videos.



Feature Extraction

We evaluated five separate feature extraction methods. For four of them, we used the numerical coordinates of the detected hand landmarks concatenated into a 1-dimensional vector as the primary feature representation. For the remaining model, we fine-tuned a mobile-optimized CNN, MobileNetV2 [51], to learn features derived from raw image sequences. We noted that the landmark-based feature representations are privacy-preserved, as they do not require the face of the participant to be shown in the given data for adequate classification.

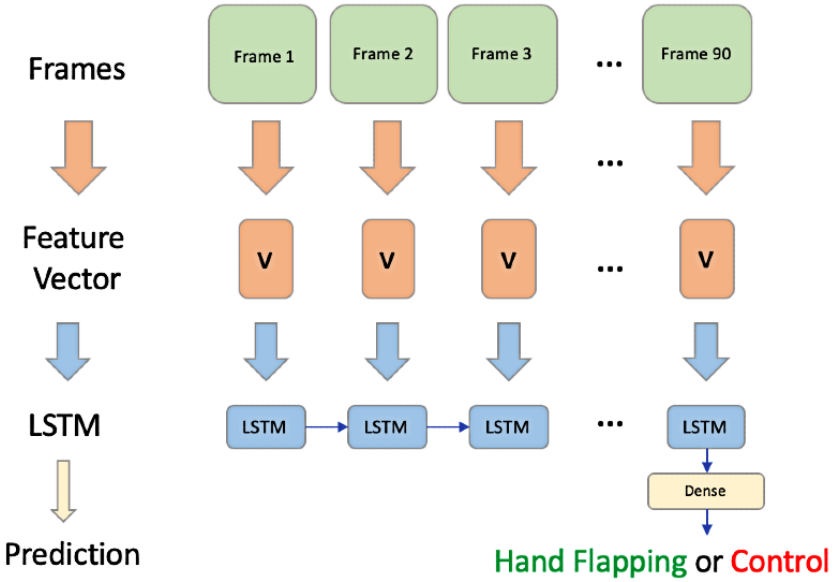
To extract the hand coordinates, we used MediaPipe, a framework hosted by Google that detects the landmarks on a person's face, hands, and body [52]. MediaPipe's hand landmark detection model provides the (x, y, z) coordinates of each of the 21 landmarks it detects on each hand. The x coordinate and y coordinate describe how far the landmark is on the horizontal and vertical dimensions, respectively. The z coordinate provides an estimation of how far the landmark is from the camera. When MediaPipe does not detect a landmark, the (x, y, z) coordinates are all set to 0 for that landmark.

The first landmark-based feature representation approach we tried used all 21 landmarks on each hand provided by MediaPipe to create the location vector fed into the LSTM. SSBD's videos mostly contain children whose detected hand landmarks are closer together due to smaller hands. This could be a problem when generalizing to older individuals with wider gaps between hand landmarks. To help the model generalize beyond hand shape, one possible solution is to use a curated subset of landmarks.

To eliminate hand shape all together, one could use only one landmark. We tried this method by using a single landmark at the base of the hand. However, because the videos in SSBD may be shaky, reliance on MediaPipe being able to detect this landmark may have led to empty features for some frames. One way to circumvent this problem is to take the mean of all the (x, y, z) coordinates of detected landmarks and use the average coordinate for each hand. We call this method the "mean landmark" approach.

We took the first 90 frames of a video and for each frame, we concatenated the feature vectors and used them as input for each timestep of an LSTM model (Figure 2). We experimented with subsets of landmarks provided by MediaPipe; we tried using all 21 landmarks, 6 landmarks (5 at each fingertip and 1 at the base of the hand), and with single landmarks. We note that the concatenated coordinates of landmarks will always form a vector that is 6 times larger than the number of landmarks used because there are 3 coordinates for a single landmark and 2 hands for which each landmark can be detected.

Figure 2. Hand flapping detection workflow. The initial 90 frames of a single video are each converted to a feature vector, consisting of either the location of coordinates as detected by MediaPipe (depicted here) or a feature vector extracted from the convolutional layers of a MobileNetV2 model. For all feature extraction methods, the resulting feature vectors are passed into an LSTM. The LSTM's output on the final timestep is fed into a multilayer perceptron layer to provide a final binary prediction. LSTM: long short-term memory.



Model Architecture

The neural network architecture we used for all experiments consisted of an LSTM layer with a 64-dimensional output. The output of the LSTM was passed into a fully connected layer with sigmoid activation to obtain a binary prediction. To minimize overfitting, we also inserted a dropout layer between the LSTM and the dense layer with a dropout rate of 30%. The landmark-based models contained nearly 3 million parameters. (Table 1). We note that the number of parameters depends on the feature approach; Table 1 shows the number of parameters based on our heaviest feature approach of using all 21 landmarks.

We experimented with other model architectures before selecting this model. We found that adding more than one LSTM or fully connected layer did not cause any notable difference in performance; thus, we removed these layers to minimize the model's capacity for overfitting. We also experimented with the output dimensionality of the LSTM; we tried 8, 16, 32, and 64. We found that using 32 and 64 performed similarly, with 64 usually performing slightly better.

Table 1. Number of parameters in the neural networks using hand landmarks as features. The two feature extraction models collectively contained 3,133,336 parameters. By contrast, MobileNetV2 feature extraction contained 2,260,546 parameters with 2 output classes.

Layer	Parameters, n
MediaPipe Hand Detector	1,757,766
MediaPipe Landmark Extractor	1,375,570
LSTM[a] (64 units)	48,896
Dropout (30%)	0
Dense	65
Total	3,182,297

[a]LSTM: long short-term memory.

Model Training

We trained all models with binary cross-entropy loss using Adam optimization [53]. We tried learning rates of 0.0005, 0.0001, 0.0005, 0.001, and 0.1, and found that in almost all cases 0.01 worked best. All models and augmentations were written using Keras [54] with a TensorFlow [55] back end run on Jupyter. No GPUs or specialized hardware were required due to the low-dimensional feature representation, and training a single model took a few minutes on a CPU with 32GB of RAM.

For all models, we trained the model until there was consistent convergence for 10 or more epochs. This resulted in 75 epochs of training across all models. After training, we reverted the model's weights to its weights for which it performed best. We used this strategy for all feature approaches.

Results

Overview

We used 5-fold cross validation to evaluate each model's average accuracy, precision, recall, and F1 score across all folds

for training and testing. However, because of our small data set, the particular arrangement of the videos in each fold substantially affected the model's performance. To minimize this effect, we ran the 5-fold cross-validation procedure 100 times, each with a different random seed, resulting in a total of 500 distinct folds. We further ensured that each fold was completely balanced in both the training and testing set (50% head banging and 50% not head banging). In all folds, there were 10 videos displaying hand flapping and 10 videos displaying head banging.

We report the mean and SD of each metric across all 500 folds as well as the area under receiver operating characteristics (AUROC). For all feature approaches, we also show the average receiver operating characteristics (ROC) curve across all folds.

All Hand Landmarks

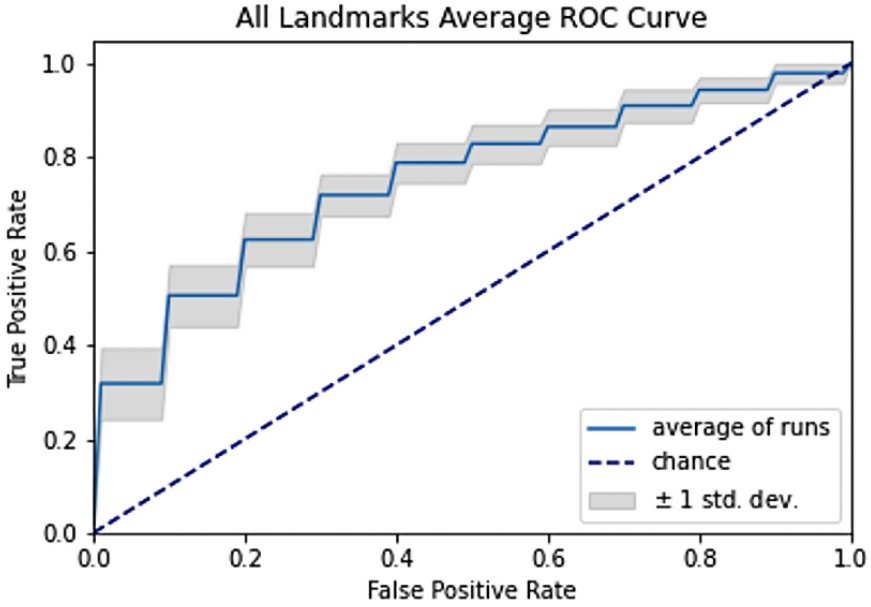
This approach used all 21 landmarks on both hands for a total of 42 unique landmarks. We show the results of this approach in Table 2. In Figure 3, we show the ROC curves of the model with and without augmentations.

When using all the landmarks, we used graphical interpolation to fill in the coordinates of missing landmarks to help reduce the effects of camera instability. However, when we tried this, we found that it often decreased accuracy and resulted in higher SDs. We therefore decided to discontinue using interpolation when evaluating the approaches described in the next section. We conjecture that the inability of MediaPipe to detect hand key points could be a salient feature for hand flapping detection, and this feature becomes obfuscated once key points are interpolated.

Table 2. Model performance for training and testing when using all hand landmarks in the feature representation.

Run type	Accuracy (SD; %)	Precision (SD; %)	Recall (SD; %)	F1 (SD; %)
Training	79.7 (1.6)	82.4 (2.67)	76.5 (3.0)	79.0 (1.7)
Testing	68.0 (2.66)	70.3 (3.6)	65.34 (5.0)	66.6 (3.35)

Figure 3. Receiver Operating Characteristics (ROC) curve across all runs when using all hand landmarks. We achieved an area under receiver operating characteristics of 0.748 (SD 0.26).



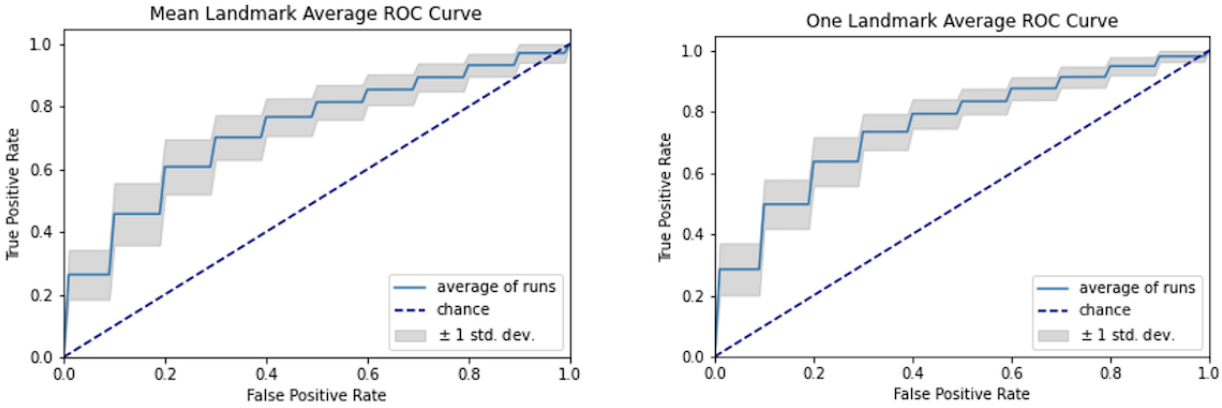
Single Hand Landmark

Here, we describe the mean and one landmark approaches, both of which relied on a single landmark on each hand as the feature representation. We show the results of both approaches, with and without augmentations, in Table 3. In Figure 4, we show the average ROC curve for both approaches.

Table 3. Model performance for mean versus single landmark feature representations with and without data augmentation.

Approach	Train/test	Accuracy (SD; %)	Precision (SD; %)	Recall (SD; %)	F1 (SD; %)
Mean landmark	Training	69.2 (4.1)	70.4 (5.3)	70.6 (7.0)	68.9 (5.12)
Mean landmark	Testing	65.5 (4.5)	66.7 (7.4)	66.9 (9.6)	64.2 (6.8)
One landmark	Training	69.2 (3.4)	70.47 (4.4)	69.71 (6.7)	68.7 (4.4)
One landmark	Testing	65.8 (4.3)	66.5 (7.5)	68.0 (6.7)	64.9 (6.5)

Figure 4. Average ROC curve for the mean (left plot) and one (right plot) landmark approach. The mean landmark approach yielded an area under receiver operating characteristics (AUROC) of 0.73 (SD 0.04), and the one landmark approach yielded an AUROC of 0.751 (SD 0.03). ROC: receiver operating characteristics.



Six Hand Landmarks

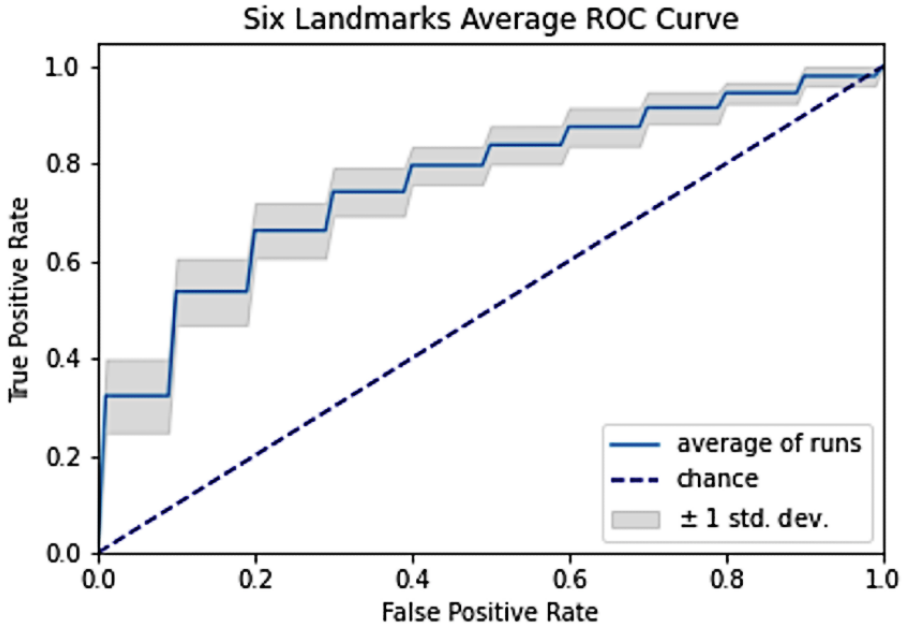
We used the six landmarks on the edges of the hands to create the location frames. We achieved an F1 score and classification accuracy of about 72.3% (Table 4). We also achieved an AUROC of 0.76 (Figure 5).

Of all of the landmark-based approaches, the six landmarks approach yielded optimal results. All of the validation metrics were higher with this approach than those previously discussed.

Table 4. Model performance in training and testing for feature representations containing six landmarks.

Run type	Accuracy (SD; %)	Precision (SD; %)	Recall (SD; %)	F1 (SD; %)
Training	76.8 (1.95)	78.7 (2.9)	74.7 (3.5)	76.2 (2.1)
Testing	69.55 (2.7)	71.7 (3.5)	67.5 (5.5)	68.3 (3.6)

Figure 5. Receiver Operating Characteristics (ROC) curve for the six landmarks approach across all runs. We achieved an area under receiver operating characteristics of 0.76 (SD 0.027) with this approach.



MobileNetV2 Model

In the approaches discussed so far, MediaPipe was consistently used as a feature extractor to bring each video frame into a lower-dimensional vector representation. Here, we substituted the MediaPipe feature extractor with MobileNetV2's [51] convolutional layers (pretrained on ImageNet [56] and fine-tuned on SSBD) as a feature extractor. As with the landmark-based approaches, this extracted vector was fed into an LSTM network to obtain a prediction for whether hand flapping was present in the video. We evaluated this model on the same 100 data sets (500 total folds), as we used for all other approaches. The ROC curve of this model is shown in Figure 6, and the metrics are detailed in Table 5.

The MobileNetV2 model achieved an accuracy and F1 score both around 85%, surpassing the performance of all the landmark-based approaches. The MobileNetV2 models also had a higher capacity to overfit, achieving near perfect accuracies in training (>99.999%), whereas all landmark-based approaches never surpassed 90% for any of the training metrics. We conjecture that this is because the MobileNet V2 model has learned both the feature extraction and discriminative steps of the supervised learning process.

Figure 6. Receiver Operating Characteristics (ROC) curve of the Mobile Net. With this method, we achieved an area under receiver operating characteristics of 0.85 (SD 0.03).

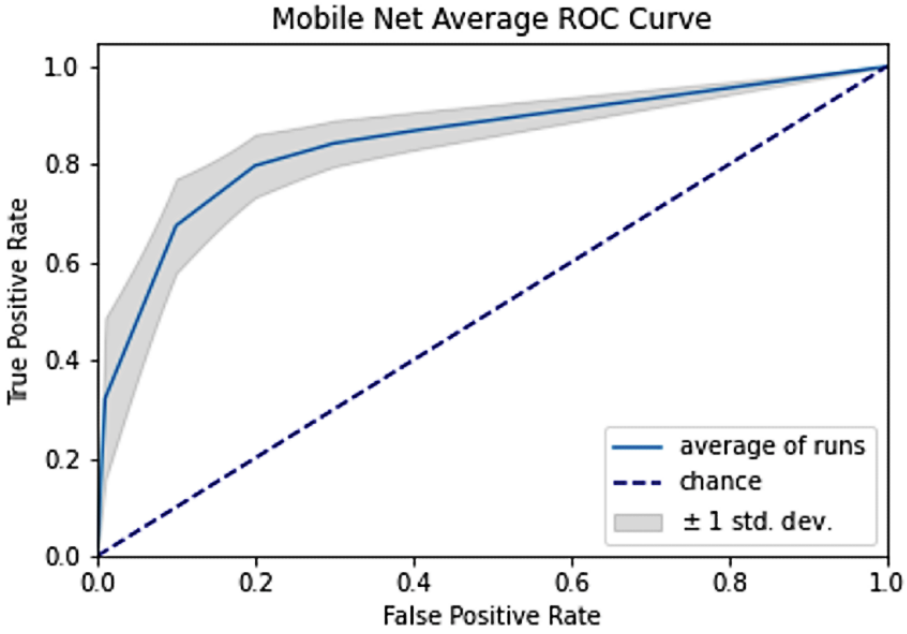


Table 5. Model performance in training and testing when using MobileNetV2 convolutional layers as the feature extractor.

Run type	Accuracy (SD; %)	Precision (SD; %)	Recall (SD; %)	F1 (SD; %)
Training	97.7 (1.0)	99.5 (0.0)	95.9 (1.7)	97.6 (1.0)
Testing	85.0 (3.14)	89.6 (4.3)	80.4 (6.0)	84.0 (3.7)

Comparison of Results

We conducted a 2-sided *t* test to determine whether the differences we observed for each approach (including the MobileNetV2 method) were statistically significant. We applied Bonferroni correction across the comparisons, deeming a *P* value <.005 as statistically significant. We show the *P* values from comparing all the approaches with each other on the 4 aforementioned metrics in Table 6.

Most of the comparisons between approaches were statistically significant after Bonferroni correction. The two single landmark approaches (mean and one landmark) were not statistically significant for any of the metrics.

Table 6. We conducted a 2-sided t test to determine whether the differences in results for each approach were statistically significant. We display *P* values for the 500 accuracy, precision, recall, and F1 values.

	All landmarks vs mean landmark (*P* value)	All landmarks vs one landmark (*P* value)	All landmarks vs six landmarks (*P* value)	All landmarks vs mobile net (*P* value)	Six landmarks vs mean landmark (*P* value)	Six landmarks vs one landmark (*P* value)	Six landmarks vs mobile net (*P* value)	Mean landmark vs one landmark (*P* value)	Mean landmark vs mobile net (*P* value)	One landmark vs mobile net (*P* value)
Accuracy	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.67	<.001	<.001
Precision	<.001	<.001	.007	<.001	<.001	<.001	<.001	.85	<.001	<.001
Recall	.15	.01	.004	<.001	.59	.66	<.001	.42	<.001	<.001
F1	.002	.02	.001	<.001	<.001	<.001	<.001	.50	<.001	<.001

Discussion

Principal Results

We explored several feature representations for lightweight hand flapping classifiers that achieved respectable performance on the SSBD. The highest-performing model used MobileNetV2 to extract features and achieved a test F1 score of 84 (SD 3.7). A model trained with all hand landmarks reached an F1 score of 66.6 (SD 3.35). A model trained with a select 6 landmarks reached an F1 score of 68.3 (SD 3.6). A model trained using a single landmark at the base of the hands reached an F1 score of 64.9 (SD 6.5).

One point of interest in this study is the trade-off between privacy-preserved solutions and performance in diagnostic machine learning tasks. While the MobileNetV2 model outperformed all the MediaPipe classifiers, the MobileNetV2 model lacks the capability to preserve the privacy of the participants, as the participants' faces were ultimately used in the data needed for classification. We expect this to be a difficulty for future research in the behavioral diagnostic space.

Limitations

The primary limitation of this approach is that without further class labels across a variety of hand-related activities and data sets, there is a probable lack of specificity in this model when generalizing to other data sets beyond the SSBD. Hands can move but not display hand flapping or self-stimulatory movement. Furthermore, stereotypic use of hands may occur in the absence of a formal autism diagnosis. Multi-class models that can distinguish hand movement patterns are required for this degree of precision. Such models cannot be built without corresponding labeled data sets, and we therefore highlight the need for the curation of data sets displaying behaviors related to developmental health care.

For this study to truly generalize, further validation is required on data sets beyond the SSBD. While the SSBD was curated with autism diagnosis in mind, the paper describing the original data set does not necessarily include children with confirmed autism diagnoses. Existing mobile therapies that collect structured videos of children with autism [16-18,40] can be used to acquire data sets to train more advanced models, and these updated models can be integrated back into the digital therapy to provide real-time feedback and adaptive experiences.

Opportunities for Future Work

There are myriad challenges and opportunities for computer vision recognition of complex social human behaviors [57], including socially motivated hand mannerisms. Additional prospects for future work include alternative feature representation and incorporation of modern architectures such as transformers and other attention-based models.

The hand movement classifier we describe here is one of a potential cocktail of classifiers that could be used in conjunction not only to extract features relevant to an autism diagnosis but also to provide insight into which particular symptoms of autism a child is exhibiting. The primary benefit of this approach is for greater explainability in medical diagnoses and a strive toward specificity in automated diagnostic efforts.

Comparison With Prior Work

Gaze Patterns

Gaze patterns often differ between autism cases and controls. Chang et al [58] found that people with autism spend more time looking at a distracting toy than a person engaging in social behavior in a movie when compared to those with typical development. This demonstrated that gaze patterns and a preference to social stimuli is an indicator of autism. Gaze patterns have been used as a feature in machine learning classifiers. Jiang et al [59] created a random forest classifier that used as an input a participant's performance in classifying emotions and other features about their gaze and face. They achieved an 86% accuracy for classifying autism with this approach. Liaquat et al [60] used CNNs [61] and LSTMs on a data set of gaze patterns and achieved a 60% accuracy on classifying autism.

Facial Expression

Another behavior feature relevant to autism detection is facial expression. Children with autism often evoke emotions differently than neurotypical peers. Volker et al [62] found that typically developing raters had more difficulty with recognizing sadness in the facial expressions of those with autism than controls. This finding was confirmed by Manfredonia et al [20] who used an automated facial recognition software to compare how easily those with autism and those who are neurotypical could express an emotion when asked. They found that people with autism had a harder time producing the correct facial expression when prompted compared to controls. People with autism typically have less facial symmetry [63]. Li et al [64] achieved an F1 score of 76% by using a CNN to extract features of facial expressions in images that were then used to classify autism. CNNs, along with recurrent neural networks [65], were also applied in Zunino et al's [66] work where videos were used to classify autism. They achieved 72% accuracy on classifying those with autism and 77% accuracy on classifying typically developing controls.

On-Body Devices

Smartwatch-based systems and sensors have been used to detect repetitive behaviors to aid intervention for people with autism. Westeyn et al [67] used a hidden Markov model to detect 7 different stimming patterns using accelerometer data. They reached a 69% accuracy with this approach. Albinali et al [68] tried using accelerometers on the wrists and torsos to detect stimming in people with autism. They achieved an accuracy of 88.6%. Sarker et al [69] used a commercially available smartwatch to collect data of adults performing stimming behaviors like head banging, hand flapping, and repetitive dropping. They used 70 features from accelerometer and gyroscope data streams to build a gradient boosting model with an accuracy of 92.6% and an F1 score of 88.1%.

Pose Estimation

Pose estimation and activity recognition have also been used to detect self-stimulatory behaviors. Vyas et al [45] retrained a

2D Mask R-CNN [46] to get the coordinates of 15 key points that were then transformed into a PoTion representation [47] and fed into a CNN model for a prediction of autism-related behavior. This approach resulted in a 72.4% classification accuracy with 72% precision and 92% recall. We note that they used a derived 8349 episodes from private videos of the Behavior Imaging company to train their model. Rajagopalan and Goecke [48] used the HDM from a video that gives the dominant motions detected to train a discriminatory model to detect self-stimulatory behaviors. On the SSBD [49], which we also used in this study, they reached an 86.6% accuracy on distinguishing head banging versus spinning behavior and a 76.3% accuracy on distinguishing head banging, spinning, and hand flapping behavior. We note that they did not train a classifier with a control class. Another effort sought to determine whether individuals with autism nod or shake their head differently than neurotypical peers. They used head rotation range and amount of rotations per minute in the yaw, pitch, and roll directions as features for the machine learning classifiers to detect autism [50]. They achieved a 92.11% accuracy from a decision tree model that used the head rotation range in the roll direction and the amount of rotations per minute in the yaw direction as features.

Acknowledgments

The study was supported in part by funds to DPW from the National Institutes of Health (1R01EB025025-01, 1R01LM013364-01, 1R21HD091500-01, 1R01LM013083); the National Science Foundation (Award 2014232); The Hartwell Foundation; Bill and Melinda Gates Foundation; Coulter Foundation; Lucile Packard Foundation; Auxiliaries Endowment; The Islamic Development Bank (ISDB) Transform Fund; the Weston Havens Foundation; and program grants from Stanford's Human Centered Artificial Intelligence Program, Precision Health and Integrated Diagnostics Center, Beckman Center, Bio-X Center, Predictives and Diagnostics Accelerator, Spectrum, Spark Program in Translational Research, MediaX, and the Wu Tsai Neurosciences Institute's Neuroscience:Translate Program. We also acknowledge generous support from David Orr, Imma Calvo, Bobby Dekesyer, and Peter Sullivan. PW would like to acknowledge support from Mr Schroeder and the Stanford Interdisciplinary Graduate Fellowship as the Schroeder Family Goldman Sachs Graduate Fellow.

Conflicts of Interest

DPW is the founder of Cognoa.com. This company is developing digital health solutions for pediatric health care. AK works as part-time consultant to Cognoa.com. All other authors declare no competing interests.

References

1. Maenner M, Shaw K, Bakian A, Bilder D, Durkin M, Esler A, et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/mmwr/volumes/70/ss/ss7011a1.htm [accessed 2022-05-31]
2. Ardhanareeswaran K, Volkmar F. Introduction: focus: autism spectrum disorders. Yale J Biol Med 2015;88:4.
3. Gordon-Lipkin E, Foster J, Peacock G. Whittling down the wait time: exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. Pediatr Clin North Am 2016 Oct;63(5):851-859 [FREE Full text] [doi: 10.1016/j.pcl.2016.06.007] [Medline: 27565363]
4. Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. Arch Gen Psychiatry 2006 Jun;63(6):694-701. [doi: 10.1001/archpsyc.63.6.694] [Medline: 16754843]
5. Sacrey LR, Bennett JA, Zwaigenbaum L. Early infant development and intervention for autism spectrum disorder. J Child Neurol 2015 Dec;30(14):1921-1929. [doi: 10.1177/0883073815601500] [Medline: 26323499]
6. Spotlight on: delay between first concern to accessing services. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/ncbddd/autism/addm-community-report/delay-to-accessing-services.html [accessed 2022-04-29]
7. Estes A, Munson J, Rogers SJ, Greenson J, Winter J, Dawson G. Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. J Am Acad Child Adolesc Psychiatry 2015 Jul;54(7):580-587 [FREE Full text] [doi: 10.1016/j.jaac.2015.04.005] [Medline: 26088663]
8. Haber N, Voss C, Daniels J, Washington P, Fazel A, Kline A, et al. A wearable social interaction aid for children with autism. arXiv Preprint posted online on April 19, 2020.
9. Daniels J, Schwartz J, Haber N, Voss C, Kline A, Fazel A, et al. 5.13 design and efficacy of a wearable device for social affective learning in children with autism. J Am Acad Child Adolesc Psychiatry 2017 Oct;56(10):S257. [doi: 10.1016/j.jaac.2017.09.296]
10. Kline A, Voss C, Washington P, Haber N, Schwartz H, Tariq Q, et al. Superpower Glass. GetMobile Mobile Computing Commun 2019 Nov 14;23(2):35-38. [doi: 10.1145/3372300.3372308]
11. Voss C, Schwartz J, Daniels J, Kline A, Haber N, Washington P, et al. Effect of wearable digital intervention for improving socialization in children with autism spectrum disorder: a randomized clinical trial. JAMA Pediatr 2019 May 01;173(5):446-454 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.0285] [Medline: 30907929]

12. Washington P, Voss C, Haber N, Tanaka S, Daniels J, Feinstein C, et al. A wearable social interaction aid for children with autism. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Presented at: CHI EA '16; May 7-12, 2016; San Jose, CA p. 2348-2354. [doi: 10.1145/2851581.2892282]
13. Daniels J, Schwartz JN, Voss C, Haber N, Fazel A, Kline A, et al. Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism. NPJ Digit Med 2018;1:32. [doi: 10.1038/s41746-018-0035-3] [Medline: 31304314]
14. Daniels J, Haber N, Voss C, Schwartz J, Tamura S, Fazel A, et al. Feasibility testing of a wearable behavioral aid for social learning in children with autism. Appl Clin Inform 2018 Jan;9(1):129-140 [FREE Full text] [doi: 10.1055/s-0038-1626727] [Medline: 29466819]
15. Voss C, Washington P, Haber N, Kline A, Daniels J, Fazel A, et al. Superpower glass: delivering unobtrusive real-time social cues in wearable systems. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. 2016 Presented at: UbiComp '16; September 12-16, 2016; Heidelburg, Germany p. 1218-1226. [doi: 10.1145/2968219.2968310]
16. Kalantarian H, Jedoui K, Washington P, Wall DP. A mobile game for automatic emotion-labeling of images. IEEE Trans Games 2020 Jun;12(2):213-218 [FREE Full text] [doi: 10.1109/tg.2018.2877325] [Medline: 32551410]
17. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Guess what?: towards understanding autism from structured video using facial affect. J Healthc Inform Res 2019;3:43-66 [FREE Full text] [doi: 10.1007/s41666-018-0034-9] [Medline: 33313475]
18. Kalantarian H, Jedoui K, Washington P, Tariq Q, Dunlap K, Schwartz J, et al. Labeling images with facial emotion and the potential for pediatric healthcare. Artif Intell Med 2019 Jul;98:77-86 [FREE Full text] [doi: 10.1016/j.artmed.2019.06.004] [Medline: 31521254]
19. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall D. A gamified mobile system for crowdsourcing video for autism research. 2018 Presented at: 2018 IEEE International Conference on Healthcare Informatics; June 4-7, 2018; New York City, NY. [doi: 10.1109/ichi.2018.00052]
20. Manfredonia J, Bangerter A, Manyakov NV, Ness S, Lewin D, Skalkin A, et al. Automatic recognition of posed facial expression of emotion in individuals with autism spectrum disorder. J Autism Dev Disord 2019 Jan;49(1):279-293. [doi: 10.1007/s10803-018-3757-9] [Medline: 30298462]
21. Chong E, Clark-Whitney E, Southerland A, Stubbs E, Miller C, Ajodan EL, et al. Detection of eye contact with deep neural networks is as accurate as human experts. Nat Commun 2020 Dec 14;11(1):6386. [doi: 10.1038/s41467-020-19712-x] [Medline: 33318484]
22. Mitsuzumi Y, Nakazawa A, Nishida T. DEEP eye contact detector: robust eye contact bid detection using convolutional neural network. 2017 Presented at: 2017 British Machine Vision Conference; 2017; London. [doi: 10.5244/c.31.59]
23. Levy S, Duda M, Haber N, Wall DP. Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism. Mol Autism 2017;8:65 [FREE Full text] [doi: 10.1186/s13229-017-0180-6] [Medline: 29270283]
24. Kosmicki JA, Sochat V, Duda M, Wall DP. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Transl Psychiatry 2015 Mar 24;5:e514. [doi: 10.1038/tp.2015.7] [Medline: 25710120]
25. Wall DP, Dally R, Luyster R, Jung J, Deluca TF. Use of artificial intelligence to shorten the behavioral diagnosis of autism. PLoS One 2012;7(8):e43855 [FREE Full text] [doi: 10.1371/journal.pone.0043855] [Medline: 22952789]
26. Tariq Q, Daniels J, Schwartz JN, Washington P, Kalantarian H, Wall DP. Mobile detection of autism through machine learning on home video: a development and prospective validation study. PLoS Med 2018 Nov;15(11):e1002705 [FREE Full text] [doi: 10.1371/journal.pmed.1002705] [Medline: 30481180]
27. Tariq Q, Fleming SL, Schwartz JN, Dunlap K, Corbin C, Washington P, et al. Detecting developmental delay and autism through machine learning models using home videos of bangladeshi children: development and validation study. J Med Internet Res 2019 Apr 24;21(4):e13822 [FREE Full text] [doi: 10.2196/13822] [Medline: 31017583]
28. Washington P, Tariq Q, Leblanc E, Chrisman B, Dunlap K, Kline A, et al. Crowdsourced feature tagging for scalable and privacy-preserved autism diagnosis. medRxiv Preprint posted online on December 17, 2020. [doi: 10.1101/2020.12.15.20248283]
29. Abbas H, Garberson F, Glover E, Wall DP. Machine learning approach for early detection of autism by combining questionnaire and home video screening. J Am Med Inform Assoc 2018 Aug 01;25(8):1000-1007 [FREE Full text] [doi: 10.1093/jamia/ocy039] [Medline: 29741630]
30. Duda M, Kosmicki JA, Wall DP. Testing the accuracy of an observation-based classifier for rapid detection of autism risk. Transl Psychiatry 2014 Aug 12;4:e424. [doi: 10.1038/tp.2014.65] [Medline: 25116834]
31. Duda M, Ma R, Haber N, Wall DP. Use of machine learning for behavioral distinction of autism and ADHD. Transl Psychiatry 2016 Mar 09;6:e732. [doi: 10.1038/tp.2015.221] [Medline: 26859815]
32. Washington P, Kalantarian H, Tariq Q, Schwartz J, Dunlap K, Chrisman B, et al. Validity of online screening for autism: crowdsourcing study comparing paid and unpaid diagnostic tasks. J Med Internet Res 2019 May 23;21(5):e13668 [FREE Full text] [doi: 10.2196/13668] [Medline: 31124463]

33. Washington P, Leblanc E, Dunlap K, Penev Y, Varma M, Jung JY, et al. Selection of trustworthy crowd workers for telemedical diagnosis of pediatric autism spectrum disorder. 2021 Presented at: Biocomputing 2021: Proceedings of the Pacific Symposium; 2021; Big Island, HI. [doi: 10.1142/9789811232701_0002]
34. Washington P, Leblanc E, Dunlap K, Penev Y, Kline A, Paskov K, et al. Precision telemedicine through crowdsourced machine learning: testing variability of crowd workers for video-based autism feature recognition. J Pers Med 2020 Aug 13;10(3):86 [FREE Full text] [doi: 10.3390/jpm10030086] [Medline: 32823538]
35. Washington P, Tariq Q, Leblanc E, Chrisman B, Dunlap K, Kline A, et al. Crowdsourced privacy-preserved feature tagging of short home videos for machine learning ASD detection. Sci Rep 2021 Apr 07;11(1):7620. [doi: 10.1038/s41598-021-87059-4] [Medline: 33828118]
36. Washington P, Leblanc E, Dunlap K, Kline A, Mutlu C, Chrisman B, et al. Crowd annotations can approximate clinical autism impressions from short home videos with privacy protections. medRxiv Preprint posted online on July 6, 2021. [doi: 10.1101/2021.07.01.21259683]
37. Washington P, Yeung S, Percha B, Tatonetti N, Liphardt J, Wall DP. Achieving trustworthy biomedical data solutions. 2020 Presented at: Biocomputing 2021: Proceedings of the Pacific Symposium; 2020; Big Island, HI. [doi: 10.1142/9789811232701_0001]
38. Washington P, Park N, Srivastava P, Voss C, Kline A, Varma M, et al. Data-driven diagnostics and the potential of mobile artificial intelligence for digital therapeutic phenotyping in computational psychiatry. Biol Psychiatry Cogn Neurosci Neuroimaging 2020 Aug;5(8):759-769 [FREE Full text] [doi: 10.1016/j.bpsc.2019.11.015] [Medline: 32085921]
39. Washington P, Kline A, Mutlu OC, Leblanc E, Hou C, Stockham N, et al. Activity recognition with moving cameras and few training examples: applications for detection of autism-related headbanging. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI EA '21; May 8-13, 2021; Yokohama, Japan p. 1-7. [doi: 10.1145/3411763.3451701]
40. Kalantarian H, Jedoui K, Dunlap K, Schwartz J, Washington P, Husic A, et al. The performance of emotion classifiers for children with parent-reported autism: quantitative feasibility study. JMIR Ment Health 2020 Apr 01;7(4):e13174 [FREE Full text] [doi: 10.2196/13174] [Medline: 32234701]
41. Washington P, Kalantarian H, Kent J, Husic A, Kline A, Leblanc E, et al. Training an emotion detection classifier using frames from a mobile therapeutic game for children with developmental disorders. arXiv Preprint posted online on December 16, 2020 [FREE Full text] [doi: 10.2196/preprints.26760]
42. Washington P, Mutlu OC, Leblanc E, Kline A, Hou C, Chrisman B, et al. Training affective computer vision models by crowdsourcing soft-target labels. arXiv Preprint posted online on January 10, 2021. [doi: 10.1007/s12559-021-09936-4]
43. Varma M, Washinton P, Chrisman B, Kline A, Leblanc E, Paskov K, et al. Identification of social engagement indicators associated with autism spectrum disorder using a game-based mobile application. medRxiv Preprint posted online on June 25, 2021. [doi: 10.1101/2021.06.20.21259187]
44. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 2000 Jun;30(3):205-223. [Medline: 11055457]
45. Vyas K, Ma R, Rezaei B, Liu S, Neubauer M, Ploetz T, et al. Recognition of atypical behavior in autism diagnosis from video using pose estimation over time. 2019 Presented at: 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing; October 13-16, 2019; Pittsburgh, PA p. 1-6. [doi: 10.1109/mlsp.2019.8918863]
46. Girdhar R, Gkioxari G, Torresani L, Paluri M, Tran D. Detect-and-track: efficient pose estimation in videos. 2018 Presented at: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 18-23, 2018; Salt Lake City, UT. [doi: 10.1109/cvpr.2018.00044]
47. Choutas V, Weinzaepfel P, Revaud J, Schmid C. PoTion: Pose MoTion Representation for Action Recognition. 2018 Presented at: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 18-23, 2018; Salt Lake City, UT. [doi: 10.1109/cvpr.2018.00734]
48. Rajagopalan SS, Goecke R. Detecting self-stimulatory behaviours for autism diagnosis. 2014 Presented at: 2014 IEEE International Conference on Image Processing; October 27-30, 2014; Paris, France p. 1470-1474. [doi: 10.1109/icip.2014.7025294]
49. Rajagopalan SS, Dhall A, Goecke R. Self-stimulatory behaviours in the wild for autism diagnosis. 2012 Presented at: 2013 IEEE International Conference on Computer Vision Workshops; December 2-8, 2013; Sydney, NSW, Australia p. 755-761. [doi: 10.1109/iccvw.2013.103]
50. Zhao Z, Zhu Z, Zhang X, Tang H, Xing J, Hu X, et al. Identifying autism with head movement features by implementing machine learning algorithms. J Autism Dev Disord 2021 Jul 11:1. [doi: 10.1007/s10803-021-05179-2] [Medline: 34250557]
51. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv Preprint posted online on April 17, 2017 [FREE Full text]
52. Lugaresi C, Tang J, Hash N, McClanahan C, Uboweja E, Hays M, et al. MediaPipe: a framework for building perception pipelines. arXiv Preprint posted online on June 14, 2019 [FREE Full text]
53. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv Preprint posted online on December 22, 2014 [FREE Full text]

54. Chollet F. Keras. 2015. URL: https://keras.io/ [accessed 2022-04-28]
55. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: large-scale machine learning on heterogeneous systems. arXiv Preprint posted online on March 14, 2016. [doi: 10.1145/3190508.3190551]
56. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. 2009 Presented at: 2009 IEEE Conference on Computer Vision and Pattern Recognition; June 20-25, 2009; Miami, FL. [doi: 10.1109/cvpr.2009.5206848]
57. Washington P, Mutlu CO, Kline A, Paskov K, Stockham NT, Chrisman B, et al. Challenges and opportunities for machine learning classification of behavior and mental state from images. arXiv Preprint posted online on January 26, 2022 [FREE Full text]
58. Chang Z, Di Martino JM, Aiello R, Baker J, Carpenter K, Compton S, et al. Computational methods to measure patterns of gaze in toddlers with autism spectrum disorder. JAMA Pediatr 2021 Aug 01;175(8):827-836 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.0530] [Medline: 33900383]
59. Jiang M, Francis SM, Srishyla D, Conelea C, Zhao Q, Jacob S. Classifying individuals with ASD through facial emotion recognition and eye-tracking. 2019 Presented at: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society; July 23-27, 2019; Berlin, Germany. [doi: 10.1109/embc.2019.8857005]
60. Liaqat S, Wu C, Duggirala PR, Cheung SS, Chuah C, Ozonoff S, et al. Predicting ASD diagnosis in children with synthetic and image-based eye gaze data. Signal Process Image Commun 2021 May;94:116198. [doi: 10.1016/j.image.2021.116198] [Medline: 33859457]
61. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation applied to handwritten zip code recognition. Neural Computation 1989 Dec;1(4):541-551. [doi: 10.1162/neco.1989.1.4.541]
62. Volker MA, Lopata C, Smith DA, Thomeer ML. Facial encoding of children with high-functioning autism spectrum disorders. Focus Autism Other Developmental Disabilities 2009 Oct 06;24(4):195-204. [doi: 10.1177/1088357609347325]
63. Guha T, Yang Z, Ramakrishna A, Grossman RB, Hedley D, Lee S, et al. On quantifying facial expression-related atypicality of children with autism spectrum disorder. 2015 Presented at: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing; April 19-24, 2015; South Brisbane, QLD, Australia. [doi: 10.1109/icassp.2015.7178080]
64. Li B, Mehta D, Aneja D, Foster C, Ventola P, Shic F, et al. A facial affect analysis system for autism spectrum disorder. 2019 Presented at: 2019 IEEE International Conference on Image Processing; September 22-25, 2019; Taipei, Taiwan. [doi: 10.1109/icip.2019.8803604]
65. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation. Defense Tech Inf 1985:318-362. [doi: 10.21236/ada164453]
66. Zunino A, Morerio P, Cavallo A, Ansuini C, Podda J, Battaglia F, et al. Video gesture analysis for autism spectrum disorder detection. 2018 Presented at: 2018 24th International Conference on Pattern Recognition; August 20-24, 2018; Beijing, China. [doi: 10.1109/icpr.2018.8545095]
67. Westeyn T, Vadas K, Bian X, Starner T, Abowd GD. Recognizing mimicked autistic self-stimulatory behaviors using HMMs. 2005 Presented at: Ninth IEEE International Symposium on Wearable Computers (ISWC'05); October 18-21, 2005; Osaka, Japan p. 164-167. [doi: 10.1109/iswc.2005.45]
68. Albinali F, Goodwin MS, Intille SS. Recognizing stereotypical motor movements in the laboratory and classroom: a case study with children on the autism spectrum. In: Proceedings of the 11th International Conference on Ubiquitous Computing. 2009 Presented at: UbiComp '09; September 30-October 3, 2009; Orlando, Florida p. 71-80. [doi: 10.1145/1620545.1620555]
69. Sarker H, Tam A, Foreman M, Fay T, Dhuliawala M, Das A. Detection of stereotypical motor movements in autism using a smartwatch-based system. AMIA Annu Symp Proc 2018;2018:952-960 [FREE Full text] [Medline: 30815138]

Abbreviations

AUROC: area under receiver operating characteristics
CNN: convolutional neural network
HDM: Histogram of Dominant Motions
LSTM: long short-term memory
PoTion: Pose Motion
R-CNN: region-based convolutional neural network
ROC: receiver operating characteristics
SSBD: Self-Stimulatory Behavior Dataset

Edited by A Mavragani; submitted 22.09.21; peer-reviewed by H Li, S You, S Nagavally; comments to author 14.10.21; revised version received 29.12.21; accepted 10.04.22; published 06.06.22.

Please cite as:
Lakkapragada A, Kline A, Mutlu OC, Paskov K, Chrisman B, Stockham N, Washington P, Wall DP
The Classification of Abnormal Hand Movement to Aid in Autism Detection: Machine Learning Study
JMIR Biomed Eng 2022;7(1):e33771
URL: https://biomedeng.jmir.org/2022/1/e33771
doi:10.2196/33771
PMID:27666281

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org