

Harvard Data Science Review

Personalized Data Science and Personalized (N-of-1) Trials: Promising Paradigms for Individualized Health Care

**Naihua Duan¹ Daniel Norman² Christopher Schmid^{3,4} Ida Sim⁵
Richard L. Kravitz⁶**

¹Department of Psychiatry, Irving Medical Center, Columbia University, New York City, New York, United States of America,

²Santa Monica Sleep Disorders Center, Los Angeles, California, United States of America,

³Center for Evidence Synthesis in Health, Brown School of Public Health, Brown University, Providence, Rhode Island, United States of America,

⁴Department of Biostatistics, Brown School of Public Health, Brown University, Providence, Rhode Island, United States of America,

⁵Division of General Internal Medicine, University of California San Francisco, San Francisco, California, United States of America,

⁶Department of Internal Medicine, University of California Davis, Sacramento, California, United States of America

Published on: Sep 08, 2022

DOI: <https://doi.org/10.1162/99608f92.8439a336>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

ABSTRACT

The term ‘data science’ usually refers to the process of extracting value from *big data* obtained from a large group of individuals. An alternative rendition, which we call *personalized data science* (Per-DS), aims to collect, analyze, and interpret *personal data* to inform *personal* decisions. This article describes the main features of Per-DS, and reviews its current state and future outlook. A Per-DS investigation is of, by, and for an individual, the Per-DS investigator, acting simultaneously as her own *investigator*, *study participant*, and *beneficiary*, and making *personalized* decisions for study design and implementation. The scope of Per-DS studies may include systematic monitoring of physiological or behavioral patterns, case-crossover studies for symptom triggers, pre-post trials for exposure–outcome relationships, and personalized (N-of-1) trials for effectiveness. Per-DS studies produce *personal knowledge* generalizable to the individual’s future self (thus benefiting herself) rather than knowledge generalizable to an external population (thus benefiting others). This endeavor requires a pivot from *data mining* or *extraction* to *data gardening*, analogous to home gardeners producing food for home consumption—the Per-DS investigator needs to ‘*cultivate the field*’ by setting goals, specifying study design, identifying necessary data elements, and assembling instruments and tools for data collection. Then, she can implement the study protocol, harvest her personal data, and *mine* the data to *extract* personal knowledge. To facilitate Per-DS studies, Per-DS investigators need support from community-based, scientific, philanthropic, business, and government entities, to develop and deploy resources such as peer forums, mobile apps, ‘virtual field guides,’ and scientific and regulatory guidance.

Keywords: case-crossover design, heterogeneity of treatment effects, monitoring, pre-post trials, self-tracking, symptom triggers

Media Summary

Data science is commonly construed as the process of extracting or mining knowledge from ‘big data’ obtained from a large group of individuals, for insights that can be used to shape clinical, corporate, or public policies. This article introduces a complementary construction: *personalized data science* (Per-DS), the scientific investigation of an individual’s own data. Each individual’s Per-DS investigation produces *personal knowledge* meant to benefit the individual herself, rather than *generalizable knowledge* meant to benefit others. The individual Per-DS investigator acts simultaneously as the *investigator*, *study participant*, and *beneficiary* of her own study. Such studies require the active involvement of the individual in study design, data collection, analysis, and interpretation—a process we call *data gardening*, analogous to home gardeners producing food for home consumption, to highlight the need to ‘cultivate the field’ in order to produce personal data from the rich terrain of daily life, to be harvested for personal knowledge to inform the individual’s personal decision. Per-DS investigations can be used to identify aberrations in an individual’s physiology or behavior; to ferret out symptom triggers; to evaluate relationships between exposures (drugs, nutrients, environmental agents, and

behaviors) and outcomes; and to compare the effectiveness of treatments. Those investigations derive knowledge directly from ‘*the patient that is me*,’ rather than relying on proxy results from ‘*patients like me*.’

Despite the appeal of Per-DS, numerous barriers threaten its uptake. Importantly, most individuals who are interested in conducting their own Per-DS investigations cannot do it on their own. They need social support from peers, resource support from public and private sectors, and support from the data science community and subject area experts to step up as ‘civil engineers’ to build the needed infrastructure and tools, such as virtual peer forums; ‘virtual field guides’ to assist individual Per-DS investigators with their specific studies; user-friendly apps to facilitate self-administered investigations; templates for study design; questionnaires and sensors for data collection; analytic algorithms; and tools to aid review and interpretation of results. These support groups will, in turn, require ethical and regulatory guidance to ensure safety and effectiveness of new Per-DS approaches and to optimize their implementation.

The prospect of bringing data science into millions of ‘*home data gardens*’ is both a daunting challenge and a tremendous opportunity. Numerous home gardeners take pride in their tasty beefsteak tomatoes, with help from peer home gardeners and reputable suppliers for seeds, fertilizer, guidebooks, and so on. Many home data gardeners might also be ready for their home data ‘tomatoes,’ with help from peer data gardeners and support from data science and health science ‘suppliers.’ If successful, these efforts could produce not only benefits for individual well-being, but also a more inclusive and less hierarchical knowledge enterprise, and a culture of evidenced-based decision-making.

1. Introduction

Data science has emerged in this century as an influential paradigm for evidence-based decision-making. The prevailing version of data science is associated with ‘big data’ obtained from a large group of individuals. This rendition of data science might be characterized as *population-based data science* (*Pop-DS*, to be pronounced POP-dee-ess),¹ which utilizes data from a large group of individuals to extract *generalizable knowledge* to inform policy decisions targeted at the population or subpopulations of interest, as well as individual decisions targeted at specific individuals of interest—further discussion on the latter is given in [Section 4.1.1](#).

An alternative to the ‘big data’ rendition of data science is *personalized data science* (*Per-DS*, to be pronounced PER-dee-ess), the scientific investigation of an individual’s own data, including the design of the data collection protocol, production of the data, and extraction of personal knowledge from the data thus produced, to inform the individual’s personal decisions.

Given our primary interest and expertise in health and wellness, we focus this article on Per-DS applied to personal health and wellness decisions,² including uptake, calibration, and discontinuation of medical treatments as well as therapeutic or preventive health behaviors such as exercise and diet. We recognize that the technologies discussed in this article for personal health care decisions can also be applied to personal

decision-making in other domains, such as food preparation, personal finance, transportation, home energy use, recreation, shopping, and musical or artistic pursuits.

Per-DS studies include (but are not limited to) personalized (N-of-1) trials, that is, multiple crossover trials of two or more interventions within a single individual, to compare the effectiveness of those interventions for that individual. In addition to those trials, other important types of Per-DS studies may include systematic monitoring of physiological or behavior patterns, case-crossover studies for symptom triggers and deterrents, and pre-post trials for exposure–outcome relationships. Further discussion of the typology of Per-DS studies is given in [Section 3](#).

We shall refer to a practitioner of Per-DS as a *Per-DS investigator*, defined as an individual who conducts a Per-DS investigation with her own data, to inform her own decisions.³ The Per-DS investigator serves multiple roles, not only as the study’s investigator, but also as the study’s solo participant,⁴ and the beneficiary who consumes the personal knowledge produced in the study to inform her health/wellness decisions going forward (Figure 1). Throughout the investigation, the self-investigator might partner with health or wellness professionals (clinician, coach/trainer, etc.) and/or data science professionals, to enhance the success of the self-study. The self-study might stimulate new interests in new self-studies to resolve new questions that emerge from the current study, motivating a continuous learning experience.

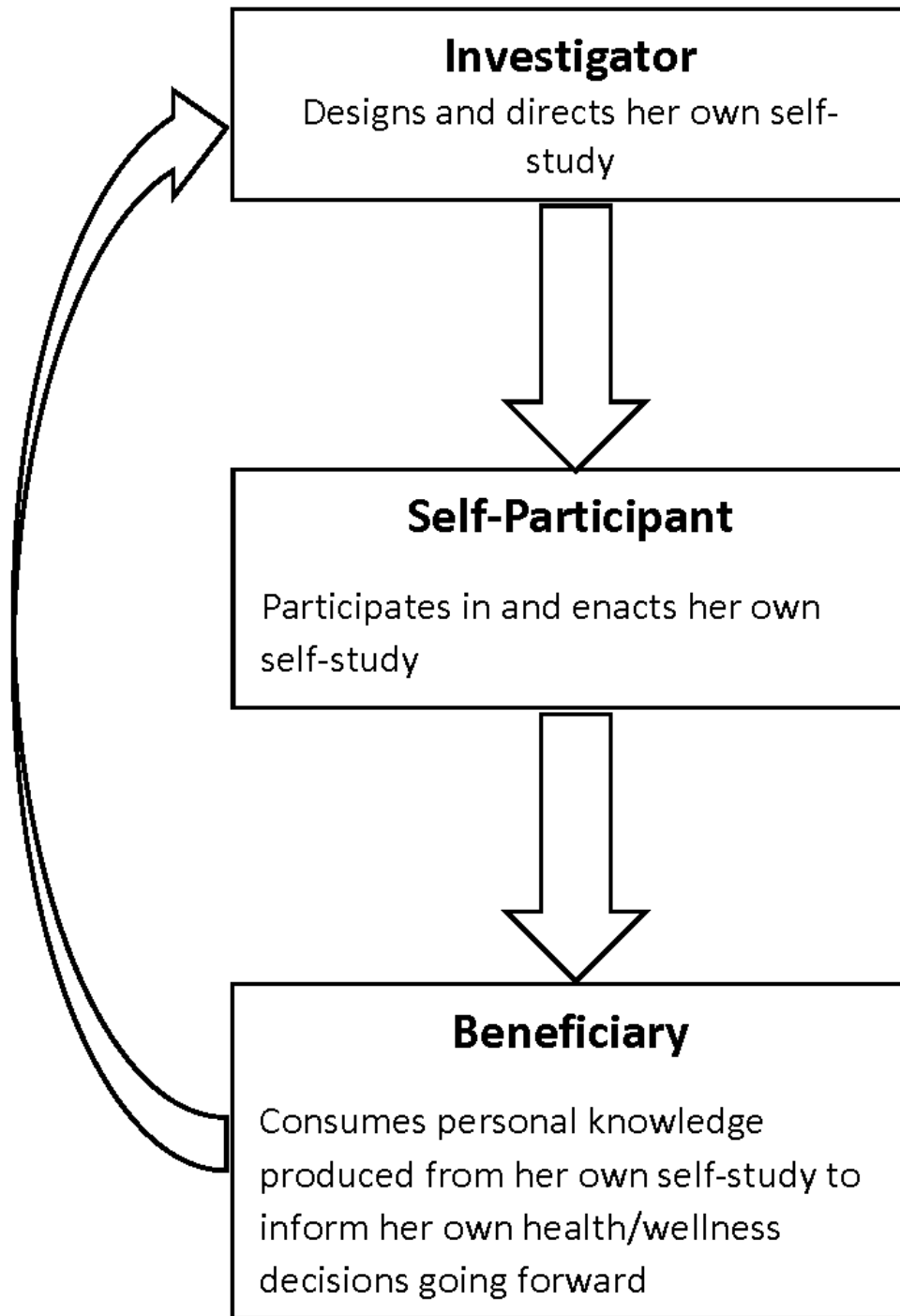


Figure 1. Per-DS investigator serving triple roles simultaneously.

Per-DS differs from Pop-DS in several important ways:

- From whom data are collected;
- Whether the study protocol is standardized or personalized;
- How knowledge is produced;

- The type of knowledge produced;
- To whom the knowledge is applied; and
- Who controls the operation?

These distinctions are summarized in Table 1 and further elucidated in [Section 4](#).

Table 1. Distinguishing features for Per-DS versus Pop-DS.

Key Dimension	Pop-DS	Per-DS
Source of data	Study sample for a target population, or ‘statistical neighbors’ for a targeted individual	Individual serving simultaneously as investigator, study participant, and beneficiary
Study protocol	Standardized	Personalized
Mode of knowledge production	Extraction (‘mining’) from large data repositories	‘Gardening’ of home-grown personal data
Nature and beneficiary of knowledge produced	Generalizable knowledge to benefit others	Personal knowledge to benefit Per-DS investigator herself
Controlling entity	Experts (scientists, policy makers, institutions)	Per-DS investigator, maybe with help from her clinician and/or data scientist

Note. Per-DS = personalized data science; Pop-DS = population-based data science.

Per-DS is an integral component of *personal science*, defined by Wolf and De Groot (2020, p.1) as “using empirical methods to pursue personal health questions,” and “consists of five activities: questioning, designing, observing, reasoning, and discovering.” Per-DS relates to personal science just as data science relates to science. Data science collaborates with subject area science to identify, collect, manage, and analyze data to advance scientific knowledge. Similarly, Per-DS collaborates with subject area *personal science* to identify, collect, manage, and analyze personal data to advance personal scientific knowledge.

Personal data for use in Per-DS investigations have also been called *small data*, described as ‘digital traces’ that individuals leave as they go about their daily activities or that individuals consciously collect using sensors (wearable and otherwise) (Estrin, 2013, 2014, 2019; Hsieh et al., 2018). In this article, we use the term ‘personal data’ instead of ‘small data,’ in part because some personal data sets can be quite large, and in part because the terminology ‘small data’ has been used with different connotations in various other contexts (e.g., Lindstrom, 2016).

To focus ideas, this article pivots on Per-DS investigations conducted by an individual (maybe with a clinician's assistance) using her own data to inform her health/wellness decisions, without sharing data with others. It is possible to extend this 'pure' Per-DS model into hybrid models, such as when a group of Per-DS investigators share data across similar Per-DS studies, or with investigations organized by researchers or health care organizations. Such hybrid models are discussed further in [Section 6.2](#).

This article describes the main features of Per-DS and reviews its current state and future outlook. The target audience for this article includes end users (patients,⁵ in some cases partnered with their clinicians) interested in conducting Per-DS investigations; clinicians, health scientists, and data scientists interested in facilitating and exploring applications of such studies; entrepreneurs viewing Per-DS infrastructure-building as a venture opportunity; and policymakers interested in new directions for health care policy (Selker et al., 2021).

2. Why *Personalized Data Science* (Per-DS)?

The rationale for Per-DS is analogous to the rationale for personalized medicine (sometimes called precision medicine) (Phillips, 2020).

While population-based medicine, largely based on the results of parallel group clinical trials, has demonstrated the efficacy of many successful treatments in groups of individuals, treatment decisions must often be personalized to accommodate individuals' unique needs, preferences, capabilities, and characteristics. However, individuals might respond differently to the same treatments, a phenomenon called heterogeneity of treatment effects (HTE) (Kravitz et al., 2004). Therefore, in many circumstances, personalized medicine tailored to the individual might outperform one-size-fits-all approaches.

With Per-DS investigations, it might also be important to personalize each investigation's *attributes* to accommodate each individual's unique needs, preferences, capabilities, and characteristics. By Per-DS 'attributes' we mean how a Per-DS investigation is designed and implemented, including what questions are selected for investigation; how data are collected, managed, and analyzed; and how the results of the investigation are interpreted and presented to the end user. Examples are given below.

2.1. The Scope of Personalization in Per-DS Studies

In traditional clinical trials, design and implementation attributes are usually standardized for all trial participants. This makes sense because the purpose of these trials is to maximize between-individual generalizability of the knowledge produced from trial participants to *other individuals*.

In contrast, between-individual generalizability is not critical for Per-DS investigations. Rather, it is important to align the design and implementation attributes for the Per-DS investigation with the needs, preferences, capabilities, and characteristics of the Per-DS investigator, to inform decisions for her own future health/wellness.

Numerous Per-DS design and implementation attributes are amenable to personalization. Some examples are shown in Table 2, drawing upon experience from the Personalized Research for Monitoring Pain Treatment (PREEMPT) Study (Barr et al., 2015; Kravitz et al., 2018).

Table 2. Personalization of Per-DS attributes considered in PREEMPT study (Barr, et al. 2015; Kravitz, et al. 2018).

Category	Attribute	Options Considered/Offered to PREEMPT Participants*	
Trial design	Pain management treatments to be compared	Same two treatments compared for all participants	*Each participant selects two treatment sets from list of suitable treatment options (e.g., NSAID, opioid, NSAID + opioid, complementary/alternative treatments)
	Duration of treatment period	Fixed for all participants	*Two options: one vs. two weeks
	Stopping rule	*Fixed	Sequential or adaptive
	Total number of treatment periods under fixed stopping rule	Fixed for all participants	*Three options: four, six, or eight treatment periods
	Washout between treatment periods	Yes, set time with neither treatment between treatment periods	*No, treatment switched immediately without gap between periods, with option for subsequent ‘analytic washout’ such as discarding outcomes at beginning of each treatment period.
	Blinding of treatment assignment	Yes	*No

Outcomes	Outcomes to be measured	Fixed for all participants	*Core outcomes fixed, option for each participant to choose additional outcomes (like neuropathic pain)
	Primary outcome for each personalized trial	Fixed for all participants	*Based on patient preference (may select an outcome different from primary outcome for overall study)
	Frequency of outcome assessments	*Daily	Multiple times per day, or ecological momentary assessment (Shiffman 2008)
Trial implementation	Use of reminders to take treatments and enter patient-reported outcomes	*Yes	No
	Frequency and timing of reminders if used	Fixed	*Specify according to patient preference
	Ability for patient to adjust or turn off reminders	*Yes	No
Data analysis	Ability for patient to access interim data and interim analysis results	*Yes	No
	Analysis plan (both interim and final)	*Predetermined	Specified according to patient preference
Presentation of results	Display format	*Graphical	Tabular or hybrid
	Details provided	*Simple point estimates	* Comprehensive, including point estimates and margin of error

* Attribute options marked with an asterisk were implemented in the PREEMPT Study. Other options were considered but not implemented.

Note. Per-DS = personalized data science.

When given the opportunity, PREEMPT study participants exercised their options to personalize their individual Per-DS studies. For example, they selected a rich variety of treatments to be evaluated:

Of the treatment categories, 31% of patients incorporated acetaminophen, 57% a nonsteroidal anti-inflammatory drug, 24% tramadol, 26% incorporated an opioid (eg, codeine, hydrocodone, or oxycodone combination product), and 48% incorporated 1 or more nonpharmacologic, complementary, or alternative treatments (eg, exercise, physical therapy, tai chi, massage, acupuncture, mindfulness meditation). (Kravitz et al., 2018, pp. 1372–1373)

Furthermore, the participants also made personalized selections for “Duration of treatment period” (28 selected “One week,” 74 selected “Two weeks”) and “Total number of treatment periods” (64 selected “Four,” 32 selected “Six,” and 6 selected “Eight.”) Apparently, some participants preferred to finish their personalized trials sooner, while some preferred to devote more time and effort, maybe hoping to achieve more reliable results.

2.2. Personalizing Results Presentations

We now elaborate on the last part of [Table 2](#) and discuss the need to personalize ‘Presentation of results,’ an important attribute for Per-DS investigations.

Research about presenting scientific information to nonscientists is quite extensive (Bonner et al., 2021; Franconeri et al., 2021; Kale et al., 2021; Kim et al., 2021). For example, a recent study represented a probability P with a set of 100 stick figures with $100 \times P$ of them shaded the same color (Marcus et al., 2022). Videos carefully explaining the meaning of the results offer another option to improve comprehension (Kravitz et al., 2018).

To inform personal decisions most effectively, Per-DS investigations should tailor presentation of findings to the preferences, educational backgrounds, and learning styles of each individual end user. For example, a visual learner with few skills in mathematics and statistics might prefer simple bar charts without measures of uncertainty, whereas someone more facile with numbers might want tables or figures with point estimates and margins of error.

Whitney et al. (2018) examined participant preference for six forms of graphical results presentations from Per-DS investigations conducted among participants with chronic musculoskeletal pain in the PREEMPT study (Kravitz et al., 2018). Each participant conducted a personalized trial comparing two pain treatment regimens. The graphs displayed how the participant fared under each treatment regimen along six outcome dimensions. The six displays ranged from simple bar charts to more complex figures showing point estimates, margins of error, and posterior probabilities of outcomes.⁶ Among 30 participants interviewed, the display of simple bar charts was rated as the “most helpful for decision making” by nine participants (30%), while a more complex graph with margins of error was rated as the most helpful by seven participants (23%).

The results of this study underscore the importance of flexibility in designing effective Per-DS investigations. Just as it would be a mistake to drown all Per-DS investigators in complex statistical information, it would be

equally wrong to provide everyone with only simple bar charts. In designing apps to support Per-DS investigations, developers should keep in mind this ‘flexibility principle.’ To avoid offering too many choices (something that can confuse consumers) (Schwartz & Ward, 2004), developers could provide a default option (e.g., simple bar charts as the default display), with other choices available for those with different preferences.

Beyond offering flexibility in the statistical details to be personalized to accommodate individual capabilities and preferences, it is also important to offer narrative text, for example, ‘it is quite likely that A is somewhat better than B,’ as an option for end users who prefer textual summaries instead of statistical summaries.

The ‘meta-design’ of Per-DS study apps, such as the attributes to be personalized, and the options to be offered, deserve further research to enhance the fit between the meta-design and individual capabilities and preferences. It will be helpful to debrief end-users to learn about the reasons for their selections. It is important for this research to be conducted with diverse populations to capture differences in socioeconomic status, cultural background, and individual capabilities and preferences.

3. Typology of Per-DS Investigations

Per-DS investigations span a wide variety of methods and purposes, including (but not limited to) systematic monitoring of physiological or behavioral patterns for risk reduction and early detection of health problems, case-crossover studies to identify symptom triggers and deterrents, pre-post trials to evaluate exposure–outcome relationships, and personalized trials to evaluate treatment effectiveness. These major categories of Per-DS investigations are shown in Table 3 and discussed in [Sections 3.1–3.4](#).

Table 3. Typology of personalized data science (Per-DS) investigations.

Purpose	Method	Required Data
Risk reduction and early detection of health problems	Systematic monitoring of physiological or behavioral indicators	Symptoms, biomarkers, clinical events; Behaviors (e.g. adherence); Treatments, exposures; Covariates
Identification of symptom triggers and deterrents	Case-crossover design based on case history	Symptoms, clinical events; Behaviors; Exposures; Covariates
Evaluation of exposure–outcome relationships	Pre-post trial	Symptoms, biomarkers, clinical events; Exposures; Covariates
Evaluation of treatment effectiveness	Personalized trial	Symptoms, biomarkers, clinical events; Treatments, exposures (assigned experimentally); Covariates

3.1. Systematic Monitoring for Risk Reduction and Early Detection

Systematic monitoring of symptoms, physiologic parameters, and behaviors may be important for the prevention and early detection of many health conditions.

Continuous glucose monitors (CGMs) have emerged in recent years as a valuable Per-DS tool for diabetes patients to monitor their blood glucose continuously to evaluate their ‘time in range,’ the percentage of time the blood glucose falls in the target range, usually targeted to be at least 70% between 70 and 180 mg/dL (American Diabetes Association, 2020). The personal knowledge gained can be helpful in finding out which types of foods and what activity level cause the individual Per-DS investigator’s blood glucose to rise and fall.

During the COVID-19 pandemic, health care workers in most settings have monitored and reported symptoms such as cough, fever, chills, shortness of breath, and gastrointestinal disturbances (Centers for Disease Control and Prevention, 2022). The knowledge gained from such systematic monitoring informs the individual’s decisions to seek medical attention and self-isolate, as well as organizational decisions about the duration of required leaves of absence.

Li et al. (2017) monitored an individual’s heart rate, oxygen saturation, skin temperature, physical activity, and radiation exposure using seven wearable biosensors collectively recording more than 250,000 measurements daily for 24 months. Among other things, the monitoring data detected abnormally elevated heart rate and skin temperature, which facilitated an early diagnosis of Lyme disease. Although the multichannel biosensors used in this study might be beyond the means of home-based Per-DS investigators, systematic monitoring using personal devices (wearable or not) for heart rate and body temperature is feasible for many.

Lee et al. (2021) demonstrated the feasibility for older men receiving chronic tamsulosin therapy to manage lower urinary track symptoms (LUTS) to monitor their urinary symptoms and medication side effects using a mobile app. Daily symptom monitoring had no adverse effects on the secondary outcomes.

Forsdyke (2015) monitored his blood pressure (BP) at home at least once daily for more than a decade, observing a summertime ‘dip’ in BP. These observations allowed for downward medication dosage adjustment during the warm summer months.

David et al. (2014) studied human microbiota using stool and saliva samples collected daily from two individuals over the course of a year, along with lifestyle data collected using a diary app to record data each day for 349 health and lifestyle variables including fitness, diet, exercise, bowel movements, mood, and illness. This study found that overall human microbial communities were stable for months, while certain life events such as enteric infection and travel to a foreign country can lead to profound changes in the microbial communities.

3.2. Case-Crossover Studies to Identify Symptom Triggers and Deterrents

Many health conditions have environmental, dietary, or behavioral triggers and deterrents, such as caffeine, alcohol, and exercise, which can provoke or mitigate bothersome symptoms such as atrial fibrillation, migraine, vertigo, asthma, functional gastrointestinal disorders, and panic attacks. Once key triggers and deterrents have been identified for an individual, she can reduce flareups by reducing exposure to the triggers, taking mitigation measures when exposed to the triggers, or by increasing exposure to the deterrents. As these associations are often individual-specific, Per-DS investigations may be particularly useful for producing such personal knowledge.

Drangsholt (2016) used case-crossover methods (Maclure, 1991; Maclure & Mittleman, 2000) to study personal triggers for paroxysmal atrial fibrillation (AF), using his own observational data. As shown in [Table 4](#), the case-crossover approach tracks occurrences of events (here, episodes of AF) and then examines an exposure period preceding the event (e.g., 24 hours) for potential triggers. A control time unit (in this case, a 24-hour period without AF) is sampled to match each case time unit. Using this method, Drangsholt learned that occurrence of his AF was associated with lack of sleep, drinking more than one glass of wine, imbibing caffeinated beverages, and public speaking. By attending to these triggers, he was able to reduce symptomatic recurrences dramatically.

Table 4. Case crossover study example.

	Case outcome (AF day)	Control outcome (no AF day)
Exposure (< 6 hours sleep)	2	6
No Exposure (6 + hours sleep)	1	19

Note. During 28-day study period, having less than 6 hours of sleep was associated with an atrial fibrillation (AF) event the next day. Odds ratio= $(A \times D) / (B \times C) = (2 \times 19) / (1 \times 6) = 6.3$.

Following the identification of triggers or deterrents using the case-crossover method, it might be useful for the Per-DS investigator to conduct further experimental studies to confirm the triggers or deterrents, using a pre-post comparison ([Section 3.3](#)) or a personalized trial ([Section 3.4](#)). Some Per-DS investigators with strong prior information for a specific trigger or deterrent might skip the observational study and proceed directly to an experimental study.

3.3. Pre-Post Trials to Evaluate Exposure–Outcome Relationships

We now turn to the use of experimental methods for evaluating the effect of interventions on outcomes. The most rudimentary experiment is arguably a pre-post trial. This approach can be used to evaluate the impact of a

specific exposure on an outcome, with the aim of informing exposure management in the future—increasing the exposure if beneficial, decreasing if harmful. The outcome may be a symptom, a dimension of subjective well-being, a physiologic parameter like blood pressure, or a biomarker like blood glucose.

The first author (ND) has a family friend who was concerned that certain restaurant food might cause blood sugar spikes, jeopardizing his efforts to maintain good diabetes control. Therefore, he regularly measured blood glucose before, immediately after, and a few hours after eating a restaurant dish. Based on the results, he stopped eating foods that were found to be associated with blood sugar spikes.

Zeevi et al. (2015) studied the postprandial glycemic response (PPGR) in an 800-person cohort of adults aged 18–70 equipped with continuous glucose monitors (CGMs). They found a high interpersonal variability in the response to standardized meals, suggesting the need to personalize dietary interventions. More broadly, CGMs can be a valuable Per-DS tool for diabetic patients to evaluate their personal responses to various types of food, to inform their dietary choice. These personalized observations may in turn drive more detailed studies designed to address *why* individual diabetic patients react differently to different diets.

The pre-post trial assumes that the outcome would have been stable if the intervention (in this case, consumption of a particular food) did not occur. The validity of this assumption needs to be evaluated carefully to avoid confounding with other changes that might have occurred during the same time period.

3.4. Personalized Trials to Evaluate Treatment Effectiveness

A Per-DS investigator might apply a personalized (N-of-1) trial to compare the effectiveness and tolerability of two treatments,⁷ perhaps to decide whether to stay with the current treatment that is working reasonably well but not perfectly, or to switch to an alternative treatment that might work better. This design assigns treatments to time periods (such as weeks), usually in blocks of two periods, with a minimum of two blocks (four treatment periods). The individual switches between the two treatments over time, for example, taking treatment A during week 1, crossing over to treatment B for week 2, then B for week 3, then A for week 4, and so on. The outcomes observed during A weeks are compared to the outcomes observed during B weeks to evaluate the comparative effectiveness for the two treatments. The results of this comparison can then inform the Per-DS investigator's future treatment decision.

This design allows the Per-DS investigator to evaluate treatment effectiveness while assessing variability across treatment periods and adjusting for time trend. For example, if there is a concern that the outcome might deteriorate slowly over time, the study can use a restricted randomization scheme, randomly choosing between the assignment sequences 'ABBA' and 'BAAB,' which are capable of mitigating the confounding arising from a linear time trend.⁸

Personalized trials are useful for chronic conditions that are stable over time.⁹ With chronicity and stability, the personal knowledge gained during the trial can inform post-trial treatment decisions. They are suitable for

treatments with rapid uptake and washout, that is, the outcome responds rapidly to treatment crossover, so that the outcome observed during A weeks indeed reflects the effect of treatment A, and likewise for B weeks. Further details about personalized trials are given in Cheung and Mitsumoto (2022), Davidson et al. (2021), Duan et al. (2013), Kravitz et al. (2014), Kravitz and Duan (2022), Schmid and Yang (2021), and other articles in this special issue.

In addition to comparing treatments, a personalized trial can also be applied to compare the effectiveness and tolerability of different doses or intensities of a treatment that is working reasonably well but might work better with enhanced calibration, such as adjusting the duration and intensity of exercise, the frequency of reminders in mobile health interventions, the strictness of time-restricted feeding, or the time and frequency of pain medication use.

Personalized trials can also be used to study symptom triggers. A Per-DS investigator might have learned from her previous case-crossover study that a specific trigger such as coffee is likely to elevate the risk for her atrial fibrillation flare-up. The candidate trigger can then be confirmed experimentally in a personalized trial, with exposure to the trigger during A weeks and no exposure during B weeks. The events observed during the A weeks are then compared with those from the B weeks to evaluate the trigger.

Marcus et al. (2022) and Kaplan et al. (2022) collaborated with the Health-eHeart network (Health eHeart Study, 2012) to recruit patients with atrial fibrillation (AF) into a randomized trial comparing the use of personalized trials testing potential triggers of their AF versus standard monitoring of symptoms. Half of the patients undertook a 6-week personalized trial with randomized 1-week intervention periods, in which they exposed themselves to the putative trigger for three periods and avoided it for three periods. The other half received usual care. At the end, the randomized groups were compared on quality-of-life changes and AF events. In addition, the triggers studied in the personalized trials were evaluated for their effects on the individuals and on groups of individuals who evaluated the same trigger.

Personalized trials can also be useful for Per-DS investigators to study the potential benefits of off-label use of a medication (Austin et al., 2021). Given the exploratory nature of such investigations, close guidance and supervision from the clinician is essential.

It is common in clinical practice to use ‘trial of therapy’ as a crude individual crossover design, with a single crossover, to inform comparative effectiveness decisions. For example, when considering whether to switch from the current treatment to a new treatment for a chronic condition, the clinician might put the patient on the new treatment, then (during the next visit) compare the outcome under the new treatment with the outcome under the original treatment, to decide whether to make the switch permanent or to abandon the new treatment and return to the original. While this simple ‘trial of therapy’ method has its merits, there are multiple challenges to its validity, including limited precision from lack of replication and possible confounding with time trend. Since the results of such clinical experiments may affect treatment choice well into the future,

investing some upfront time and resources to improve the performance of the trial can be worthwhile. In particular, personalized trials build on the single-crossover ‘trial of therapy’ by performing multiple crossovers to provide more precise and valid personal knowledge for Per-DS investigations.

4. Distinguishing Features of Personalized Data Science (Per-DS)

We now provide further discussion on important distinctions between Per-DS and Pop-DS as summarized previously in [Table 1](#).

4.1. Source of Data

While Per-DS utilizes *personal data* obtained from the Per-DS investigator’s own sources—self-reports or personal instruments, Pop-DS usually utilizes *big data* obtained from a large number of individuals not including the Per-DS investigator.

The development of Pop-DS into a major enterprise is fueled in large part by the rapid expansion of information technology that generates huge data sets needing advanced technologies to comprehend (Chen et al., 2012). The emergence of Per-DS also results from the expansion of technologies that enable individuals to access *personal data* that used to be inaccessible and incomprehensible to them.

Pop-DS is often applied to describe the population from which the big data were obtained in order to inform population-level decisions. Those objectives are distinct from the objectives of Per-DS to inform personal decisions.

Another type of Pop-DS application that is akin to Per-DS is the use of Pop-DS to make individual predictions, serving the same objective as Per-DS but using different sources of data. We discuss individual prediction for Pop-DS and Per-DS in [Sections 4.1.1](#) and [4.1.2](#), respectively, then discuss in [Section 4.1.3](#) the important distinction between the function of Per-DS to learn from ‘*the patient that is me*’ versus the function of Pop-DS to learn from ‘*patients like me*,’ and the important role of individual idiosyncrasy in this context.

4.1.1. Individual Prediction With Pop-DS

An important application of Pop-DS is to make individual predictions based on ‘reference class forecasting,’ where decisions for an index individual are influenced by results derived from a suitably specified reference class of ‘statistical neighbors’ who are similar to the index individual in terms of prognostic characteristics, known effect modifiers, susceptibility to adverse effects of treatment, preferences for alternative outcomes, and so on (Kent & Hayward, 2007; Kent et al., 2018; Kent, Paulus, et al., 2020; Kent, van Klaveren, 2020; Li & Meng, 2021; Liu & Meng, 2016; Meng, 2014, 2018; Schandelmaier et al., 2020). Meng (2014) proposed a multiple resolution framework to achieve the appropriate trade-off between bias and variance. We can narrow the lens to zoom in on individuals who are close statistical matches to the index individual and obtain more valid but less precise predictions for the index individual. Alternatively, we can widen the lens to incorporate

data from a larger, less well-matched reference class, and we get more precise but less valid predictions (Meng, 2018).

4.1.2. Individual Prediction With Per-DS

Per-DS investigation provides an individual prediction with the smallest possible ‘reference class,’ namely, the index individual herself, and the highest resolution possible—infinite resolution in terms of Meng (2014, 2018). Therefore it might be reasonable to argue that the results of a Per-DS investigation are free from the bias due to between-individual extrapolation that afflicts the Pop-DS approach to individual prediction.

Nevertheless, Per-DS investigations are not free from bias-variance trade-off, as the amount of data available from a Per-DS investigation is usually limited. Therefore, it might still be worthwhile to ‘borrow from strength’ by combining Per-DS data with Pop-DS data to provide higher precision, while tolerating some bias due to between-individual extrapolation in the Pop-DS part of the combined data (Schmid & Yang, 2022; Zucker et al., 1997, 2010).

4.1.3. ‘*The Patient That Is Me*’ Versus ‘*Patients Like Me*’—Important Role of Individual Idiosyncrasy in Individual Prediction

Beyond the bias–variance trade-off, an important feature for Per-DS that distinguishes it from Pop-DS is the important role of individual idiosyncrasy in the context of individual prediction.

Even at very high resolution, or with an infinitesimally small reference class, Pop-DS data may still manifest some variation across individuals despite their near-identical appearance. If the residual variance does not approach zero as the resolution increases toward infinity, or as the reference class shrinks toward the index individual, it will be reasonable to infer that there is a component of individual idiosyncrasy that distinguishes each individual from her closest neighbors, that is, the individual cannot be predicted deterministically from observable data. Meng (2014, p. 542) referred to the variance for individual idiosyncrasy as the “intrinsic variance,” and noted that whether the intrinsic variance can be assumed to be zero or not “reflects whether we believe the world is fundamentally stochastic or appears to be stochastic because of our human limitation in learning every mechanism responsible for variations.”

Individual idiosyncrasy for the index individual cannot be observed in Pop-DS and needs to be considered part of the noise (unless assumed to be zero). For Per-DS, individual idiosyncrasy is observed in the index individual’s own Per-DS investigation. In statistical terms, individual idiosyncrasy is a random effect for Pop-DS even with infinite resolution and an infinitesimally small reference class, and a fixed effect instead for Per-DS.

Importantly, when the index individual is not part of the big data for the Pop-DS investigation, that is, when the Pop-DS investigation ‘did not include a subject replicating my description exactly,’ the best that can be accomplished with Pop-DS is to approximate the index individual with ‘*patients like me*,’ which treats the

index individual's idiosyncrasy as unresolved noise in the intrinsic variance, leaving a gap between '*patients like me*' and the index individual. On the other hand, the index individual's own Per-DS investigation does include the index individual herself. Therefore, while a Pop-DS investigation tells the index individual what happens to '*patients like me*,' a Per-DS investigation tells the index individual what happens to the '*the patient that is me*,' without unresolved noise in the intrinsic variance. To the extent that intrinsic variance might be present, a Per-DS investigation based on the '*the patient that is me*' provides information on individual idiosyncrasy that is unique to the index individual and is not available from Pop-DS investigations based on '*patients like me*.'¹⁰

4.2. Personalization in Per-DS Versus Standardization in Pop-DS

Per-DS investigations usually need to be personalized in order to accommodate individual investigators' unique needs, preferences, capabilities, and characteristics, as was discussed in [Section 2](#).

Pop-DS investigations are usually standardized, sometimes by design, when a single data service agency procures the entire data set, sometimes by negotiation, when collaborating entities make efforts to standardize the study protocol to facilitate data sharing (Schmid et al., 2003). Discrepancies in study protocol across collaborating entities, sometimes due to individual entities' unique needs, capabilities, and preferences to retain some autonomy over their portion of the study, might be viewed as undesirable defects in the overall study design.

4.3. Data Gardening for Per-DS Versus Data Mining/Extraction for Pop-DS

Data science is usually characterized as *the study of extracting value from data* (Columbia University Data Science Institute, 2020; Dominici & Parkes, 2021; Wing, 2019). Data science in general, and Pop-DS in particular, usually emphasizes more on data analytics, machine learning, and data mining, and less on study design and data collection. The action word '*extracting*' in the characterization of data science tends to reinforce the metaphor of '*mining*' as the governing paradigm for data science.

For Pop-DS, the extracting/mining metaphor might be appropriate when the data set has already been obtained from programmatic functions and is ready to be extracted/mined. In those situations, the data scientist might not have much control over how the data collection process was planned and implemented, although the data scientist might still need to deal with imperfections in the data collected, such as missing data, selective treatment assignment, and so on.

For Per-DS, oftentimes the data set needed does not yet exist and needs to be produced, before any analysis, mining, or extraction can begin. Therefore, the Per-DS investigator usually has the responsibility, as well as the opportunity, of starting from scratch, creating and implementing the data collection protocol, tasks that might be absent from some Pop-DS investigations.

Therefore, at least for Per-DS, a more appropriate metaphor would be *data gardening*,¹¹ analogous to home gardeners producing food for home consumption, rather than *extracting* or *mining*. The Per-DS investigator needs to first cultivate the field in order to ‘grow’ the data needed; then she will be able to harvest the data and extract the knowledge to inform her decisions. The field might be barren until cultivated. The cultivation might entail identifying the goal for the investigation, designing the study (when and how to deliver the intervention, when and how to collect the data), identifying the data elements needed, identifying and acquiring the sensors/devices/self-report instruments needed to generate/collect the data, identifying and acquiring data acquisition and analysis tools, and so on.

Under the gardening metaphor, data are not waiting passively to be extracted. Rather, data are organic and require careful nurturing to yield their informational bounty. To garden effectively, Per-DS investigators need access to sensors, instruments and infrastructure. The needed support might be accessed through support groups, commercial suppliers, or virtual extension programs supported by the data science and health science communities, analogous to the agricultural Cooperative Extension System ([Section 5.2.2.2](#)).

To summarize, *data mining* might work well for Pop-DS, but the metaphor of *data gardening* fits better for Per-DS. Mining is often the prevailing modus operandi for Pop-DS but usually does not apply to Per-DS. Gardening more aptly describes what Per-DS studies need in the way of planning, preparation, design, and collection processes, and their nature of being home-grown for home consumption. Further discussions on these metaphors for data science in general, including both Per-DS and Pop-DS, are given in [Appendix A](#), “[Should Data Scientists Be Portrayed as Gardeners/Farmers or Miners/Extractors?](#)”

4.4. Personal Knowledge From Per-DS Versus Generalizable Knowledge From Pop-DS

Pop-DS usually produces *generalizable knowledge* to be applied from one group of individuals to another. For parallel group trials, this usually entails applying the knowledge gained from trial participants to other individuals, such as future patients.

The individual conducting her own Per-DS investigation takes on three important roles simultaneously: as the *investigator*, the *study participant*, and the *beneficiary* who consumes the personal knowledge produced (Figure 1). Per-DS investigations produce *personal knowledge* that is generalized from the individual during the investigation to the same individual after the investigation.

In summary, Pop-DS generalizes across individuals (from one set of individuals to another), whereas Per-DS generalizes over time within the same individual.

4.4.1. Is Blinding Needed When Generalizing to My Future Self?

The distinction between within-individual generalizability for Per-DS and between-individual generalizability for Pop-DS has important implications on whether to use blinding in treatment studies.

Blinding is used widely in efficacy studies to suppress nonspecific effects or placebo effects.

Consider a study that aims to compare the biological efficacy for two existing medications, compound A available commercially as oval pills, and compound B available commercially as round pills. Without blinding, the effect observed combines the difference in the biological efficacy between the two compounds and the difference in the nonspecific effect between the two pill shapes. An unblinded study using existing medications is vulnerable to the confounding between the compounds and the pill shapes. Such a study might be biased for the biological efficacy of interest. Under this scenario, it can be important to blind the compounds by making the two medications into the same shape, to break the confounding, so as to provide an unbiased estimate for the biological efficacy.

For Per-DS investigations, the same individual might manifest the same nonspecific effect consistently during and after the trial, therefore the nonspecific effect should be included in the study's primary estimand, instead of suppressed as a bias. Consider a Per-DS investigator with the belief, stable over time, that oval pills work better for her than round pills. Under this assumption, her nonspecific effect favoring oval-shaped pills manifested during the trial is generalizable to herself after the trial. Therefore, it is not necessary, but actually undesirable, to ask this Per-DS investigator to blind herself to the shape of the pill. Doing so would eliminate a relevant nonspecific effect in the study, thus making the study estimate a poor approximation for the future.

Per-DS investigations to inform personal decisions should be interpreted as pragmatic studies for the evaluation of existing treatments viewed as holistic bundles, as discussed in Schwartz and Lellouch (1967, 2009). For example, assume that a Per-DS investigator wants to compare her current treatment with compound A available commercially as oval pills, versus a new treatment with compound B available commercially as round pills. The pragmatic comparison between the two treatments without blinding combines the difference in the biological efficacy between compound A versus compound B, and the difference in the nonspecific effect between oval pills versus round pills. Both components of this combination need to be captured in order to inform how the individual should choose her treatment going forward.

An efficacy trial that blinds the treatments, say, by making compound B into custom-made oval pills, captures the biological efficacy but not the nonspecific effect, therefore is biased for the pragmatic effect for this individual. After the trial, the individual can only choose between the oval-A treatment bundle versus the round-B treatment bundle. Because the custom-made oval-B treatment bundle used during the efficacy trial is not commercially available, it is not a feasible treatment option for the individual's treatment decision going forward.

4.4.2. Human Subjects Protection Regulations in Per-DS Studies

Another important implication of the within-individual generalizability for Per-DS is the weakening of the rationale for applying Human Subjects Protection regulations with the federal OHRP and local institutional review boards, which define human subject research as follows:

Research means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge. (Protection of Human Subjects, 2022)

Such regulations are usually applicable to Pop-DS investigations because the generalizable knowledge produced from those investigations matches the OHRP definition for research. As a result, it is necessary to protect study participants who contribute data to Pop-DS investigation for the benefits of other individuals.

Personal knowledge produced in Per-DS investigations, on the other hand, is meant to serve the individual's own needs, without direct application to others.¹² Therefore, the production and consumption of personal knowledge should not be considered human subjects research, and thus human subjects protection regulations should not apply (Punja et al., 2014).¹³

The exemption from human subjects protection regulations does not mean Per-DS investigations should be conducted laissez faire without guidance. Further discussions on clinician support to ensure safety for Per-DS studies are given in [Section 5.2.1](#).

4.5. Individual Control for Per-DS versus Expert Control for Pop-DS

The fifth important distinction between Per-DS and Pop-DS is the question ‘who controls the investigation, as the ultimate decision maker on design and implementation?’ Is authority for decisions on design and implementation of the investigation held by experts, or devolved to the end users? More colloquially, who holds the steering wheel?¹⁴

For Pop-DS investigations, the steering wheel is usually held by experts to serve the needs of policymakers, health care organizations, commercial entities, or academic researchers. These experts are presumably acting at least in part on behalf of the individual participants but are also accountable to the scientific community, journal editors, institutions, careers and reputations, and future generations of consumers and patients (Kravitz, 1990). The individuals participating in Pop-DS investigations usually have a passive role, such as complying with study protocols, with little or no voice in the direction of the study, other than the possibility of refusing to participate or dropping out. Many study consent forms describe study benefits and risks to participants, but these are often outside the control of the participants. Under the Pop-DS paradigm, power is distributed asymmetrically and studies are organized hierarchically, with decisions made at the top promulgated to participants at the bottom.

Per-DS investigators conduct their own investigations to inform their own personal decisions. Under this paradigm, decision-making authority should arguably be held by the Per-DS investigator herself, with suitable guidance from professionals such as her clinicians or coaches.

Per-DS constitutes both a great opportunity, and a major challenge, to clinician–patient relationships. Per-DS can be a valuable tool for clinicians to deliver enhanced care to their patients, at least to those who are

interested in conducting Per-DS investigations in partnership with their clinicians. On the other hand, some clinicians might see Per-DS as a time-sink, and a threat to their authority.

Video training could help enhance patients' self-efficacy in taking an active role in partnership with their clinicians to conduct Per-DS studies. Patient participation research networks such as Health-eHeart and the ICN registry (ImproveCareNow, 2020) might also stimulate active patient roles in Per-DS studies.

Some clinician–patient teams might agree on the need for a specific Per-DS investigation but opt for the clinician to act on the patient's behalf to design and implement the investigation. The patient might be intimidated by the overwhelming responsibility to 'hold the steering wheel' for the investigation, and thus might prefer to delegate this responsibility to her clinician. These investigations can be valid Per-DS studies as long as the clinician acts as the patient's agent to effectuate the patient's needs and preferences, instead of acting as the principal taking control of the study.

The democratic, flat structure of Per-DS is very different from the 'top-down' (authoritarian, pyramidal) organization of Pop-DS. Per-DS investigators are self-empowered with autonomy to make decisions regarding their own studies. The role of experts (e.g., a virtual extension program discussed in [Section 5.2.2.2](#)) is to support Per-DS investigators as helpers and counselors, not as decision makers holding the steering wheel. Further discussions on the potential social and political implications of the 'bottom-up' organization for Per-DS are given in [Section 6.3](#).

5. The Viability of Per-DS: Reach, Safety, and Effectiveness

The viability of Per-DS depends importantly on two questions:

- How large is the potential 'market' for Per-DS investigations? Are there enough prospective Per-DS investigators to motivate investments from public and private sources to develop support infrastructures?
- Can Per-DS investigations be conducted safely and effectively?

To address the first question, we evaluate available evidence and make additional speculative calculations. To address the second question, we consider the extant literature and imagine how safety, effectiveness, and value of Per-DS studies could be enhanced through active engagement of clinicians, scientific content experts, and data scientists.

5.1. Scoping Out the Per-DS 'Market'

Evidence for the existence of a viable Per-DS market includes personal anecdotes, several surveys, and inferences from several completed Per-DS studies.

The first author (ND) has conducted his own Per-DS studies to inform his personal decisions with his positive airway pressure (PAP) therapy for sleep apnea, including systematic monitoring for normalcy and aberrations,

and personalized trials to evaluate his PAP therapy configurations. Further discussions are given in [Appendix B, “PAP Therapy for Sleep Apnea: Fertile Field for Personalized Data Science \(Per-DS\)?”](#)

As discussed in [Section 3.3](#), ND’s family friend with diabetes conducted a series of pre-post Per-DS investigations on the impact of restaurant food on his blood glucose to inform his future dietary decisions. In addition, two of ND’s relatives with multiple prescriptions for diabetes, hypertension, and hypercholesterolemia have conducted semiformal Per-DS investigations regularly for a number of years (without ND’s input or influence) to investigate the consequences of skipping a dose of one of their medications once in a while, driven by their preference to take less medication to the extent possible. Some of their investigations utilized measured biomarkers such as blood pressure and blood glucose as their outcomes. Some relied on self-sensed outcomes such as dizziness. Favorable outcomes (little or no rise in blood pressure or blood glucose, absence of dizziness, etc.) encouraged them to conduct another similar trial soon; unfavorable outcomes encouraged them to delay the next trial to be spaced out further in time. Although neither of them devised and followed a systematic decision protocol, this informal protocol is akin to the play-the-winner adaptive design for clinical trials (Wei & Durham, 1978; Yao & Wei, 1996; Zelen, 1969).

A large number of lay Per-DS investigators use smart watches or wearable fitness trackers to track their physical activity and other health indicators to inform their health behavior decisions. For example, Vogels (2020) reported that 21% of adults in the United States reported using those devices in a survey conducted June 3–17, 2019. The use of those Per-DS devices can motivate people to exercise more (Cadmus-Bertram et al., 2015), potentially reducing their risk for cardiac disease, cancer, and fractures (Harris et al., 2019), and lowering their risk of all-cause mortality (Saint-Maurice et al., 2020).

Continuous glucose monitors have been used by 30% of the patients with Type 2 diabetes in the United States; their use is expected to rise in the years to come (Kelly, 2021). Those devices can serve as valuable Per-DS tools for Per-DS investigators to conduct Per-DS studies to improve their diabetic health, including systematic monitoring of their time-in-range for their blood glucose ([Section 3.1](#)), pre-post trials to evaluate impact of various types of food on their blood glucose ([Section 3.3](#)), and so on.

Another example of consumer demand for Per-DS is [Quantified Self](#) (n.d.), “an international community of users and makers of self-tracking tools who share an interest in ‘self-knowledge through numbers.’” The participants share their experience in a rich variety of Per-DS investigations, including self-tracking/monitoring, identification of symptom triggers, and self-experimentation, including personalized trials. There are approximately 70 local chapters, some with thousands of members.

Many sleep apnea patients are interested in learning about their condition and taking an active role in their PAP therapy. Some data scientists have developed software, free to consumers, to access and analyze the data recorded on their PAP devices, such as SleepyHead (2018) and Open Source CPAP Analysis Reporter (OSCAR) (Apnea Board, 2019). Those tools are well known among sleep apnea patients, and used regularly by

many. There are extensive discussions among patients in bulletin boards, for example, Apnea Board (2011), on practical issues such as comparison between nasal masks versus nasal pillows, getting better control of air leakage, and so on. Many of those patients have explored informal trial-and-error Per-DS investigations.

Beyond those examples, some focus groups and surveys suggested substantial consumer interest in addressing ‘what works for them’ (Derby et al., 2021; Kravitz et al., 2009). However, the available research also suggests that consumer enthusiasm diminishes rapidly as the expected burdens of participation rise. For example, Moise et al. (2018) reported that consumers want trials that are low in cost and short in duration.

Despite these limitations, the potential market for Per-DS investigations might be very large. We provide speculative calculations in two illustrative examples below.

As our first example, the Personalized Research for Monitoring Pain Treatment (PREEMPT) Study (Kravitz et al., 2018) recruited patients with chronic musculoskeletal pain and randomized them to be offered either usual care or a patient-centered personalized (N-of-1) trial. The study screened 1,092 patients and identified 360 (33.0%) as being eligible. Among those screened to be eligible and invited to participate in the study, 89.4% (=322/360) expressed an interest, and 59.7% (=215/360) actually enrolled, representing a high level of interest in personalized trials, even though they were only offered a 50% chance to be assigned to the personalized trial arm. Taking the 33.0% eligibility rate and the 59.7% enrollment rate as proxies for the U.S. population of chronic pain patients interested in some form of Per-DS investigation, and assuming that approximately 9.8 million Americans suffer from functionally important chronic musculoskeletal pain (Dahlhamer et al., 2018),¹⁵ the level of demand for Per-DS within this population could be as high as 1.9 million U.S. adults.¹⁶

As another example, one of us (DN), an experienced sleep medicine physician, estimates that approximately 40% of sleep apnea patients in his clinical practice would be interested in some form of personal investigation using their own data if clinical and technical support were available. Assuming approximately 15 million adults in the United States with moderate-to-severe sleep apnea (Howden & Meyer, 2011; Peppard et al., 2013),¹⁷ there are up to six million prospective Per-DS investigators among them.

Similarly sized markets might exist among patients with other common chronic conditions such as migraines, hypertension, diabetes, asthma, mood disorders, and so on.

5.2. Per-DS Needs Support From Health Care Professionals and Scientific Communities

So far, we have considered Per-DS investigations generically, without much consideration of safety and effectiveness. These studies can be portrayed along a spectrum of rigor, to include:

- Highly informal observations:
 - I feel better when I take vitamin C.

- Dulcolax does not work for me.
- Semi-formal search for triggers:
 - What triggers my vertigo?
- Semiformal quasi-experimental investigations:
 - My blood glucose shot up after I ate that restaurant dish.
- Formal observational study for triggers:
 - What triggers my atrial fibrillation, according to my case-crossover study?
- Formal experimental study for triggers:
 - Does alcohol trigger my atrial fibrillation, according to my personalized trial?
- Formal experimental studies for comparative effectiveness:
 - Do I sleep more comfortably with this nasal mask combined with chinstrap, than with that oronasal mask, according to my personalized trial?

Some of those self-investigations might be safe and effective, some might be ineffective, and some might even pose a risk of unintended harm to the Per-DS investigators.

It is therefore important for health care professionals and the scientific communities (both health/clinical science and data science) to support Per-DS investigators in a manner that maximizes the safety and effectiveness of their investigations.

There is a wide range of support needed for Per-DS investigations:

- Study design
 - Elimination of unsafe investigations
 - Formulation of pragmatic and effective designs
- Data collection
 - Sensing and information technologies to acquire data needed for the study
- Data management
- Data visualization and analysis
- Results interpretation

For example, many diabetic patients monitor their blood glucose regularly, using either finger stick devices or CGM devices. Their monitoring efforts will be more effective if combined with diaries of medications, activities, and other exposures to help interpret the glucose data observed. Furthermore, it is also important to take into consideration possible long-term time trends, as diabetes tends to worsen over time. Clinical input as well as data science input will enrich the value of such monitoring efforts.

Numerous apps are currently available to support diabetes patients' Per-DS investigations. It will be helpful to have apps with the ability to support personalized trials to assess how diet and exercise affect glucose levels;

these wield considerable promise but will also need careful evaluation. Unfortunately, most existing apps have not been evaluated for their usability, safety, or effectiveness (Henson et al., 2019; Lagan et al., 2020; Mohr et al., 2013, 2015). This makes it challenging for patients and their health care professionals to make informed selections. The downside potential of unfettered Per-DS app development with scant evidence-based evaluation and oversight should concern the Per-DS community and policymakers alike.

There are several possible lines of defense for safety. First comes the individual's health care professionals (primary care physicians, other clinicians, dietitians, fitness coaches, etc.), who would be in a position to advise the individual on what parameters might be measured and what treatments might be tried, given the individual's unique set of health conditions and risks ([Section 5.2.1](#)).

Beyond direct support from health care professionals, it is also important for the data science and health science communities to work proactively to develop and deploy infrastructure support for Per-DS, with built-in safety and rigor ([Section 5.2.2.2](#)).

5.2.1. Per-DS Needs Support From Clinicians

There is an important need for clinician involvement (consultation, guidance, oversight, etc.) to ensure safety for the Per-DS investigations and maybe also to enhance effectiveness of the investigations, as shown in Table 5.

Table 5. Clinician involvement needed for Per-DS investigations.

Intervention Investigated	Type of Per-DS Investigation			
	Systematic Monitoring	Case-Crossover Study for Trigger Finding	Pre-Post Trial for Exposure-Outcome Relationships	Personalized Trial for Treatment Effectiveness
Routine exercise	None	Minimum to Moderate	Minimum	Minimum
Common diet	Minimum		Low	Low
OTC medication	Low		Moderate	Moderate
Prescription medication	Moderate		High	High

Note. Per-DS = personalized data science.

Systematic monitoring based on passive observations poses low risk, therefore requires less clinician involvement than other types of investigations that might entail somewhat higher risk due to active

manipulation of the intervention. Among Per-DS investigations that involve an ‘intervention’—exercise, diet, and medication—we rank the need for clinician involvement as the lowest for routine exercise such as walking, jogging, and swimming; and slightly higher for diet due to the possible risk of adverse health outcomes. Investigations involving OTC medication have slightly higher need for clinician involvement due to the possibility of adverse effects or drug–drug interactions. Prescription medications have the highest need for clinician guidance.

Case-crossover studies for trigger finding does not involve an explicit ‘intervention.’ Rather, it is an outcome looking for an ‘intervention’ that triggers the outcome. Since it is based on passive observations and does not involve any active manipulations, we rate the need for clinician involvement as minimum/moderate, depending on the nature of the outcome. For high-risk outcomes such as atrial fibrillation, clinician involvement is more critical.¹⁸

Per-DS studies that entail active manipulation of interventions, such as pre-post trials and personalized trials, require more involvement from clinicians. Again, the level of clinician involvement needed varies with the risk level for the intervention, ranging from minimum for routine exercises to high for prescription medication.

It is important for clinicians to set the safety boundary for Per-DS investigations. Some high-risk interventions should be proscribed from self-experimentation, even with clinician supervision. These include taking new medications not yet shown to be safe, changing parameters on one’s pacemaker, or modifying the current running through one’s deep brain stimulator (DBS). (At the same time, passive Per-DS studies that entail systematic monitoring with pacemakers or DBS’s might be safe to conduct, with clinician involvement, to inform patients and their clinicians on the performance of the devices.)

5.2.2. Per-DS Needs Support From Data Science and Health Science Communities

Beyond support from clinicians, it is also important for the data science and health science communities to actively support Per-DS, both to help this nascent enterprise realize its potential, and to ensure the safety and effectiveness of the Per-DS investigations. The agenda for the data and health science communities includes: 1) assessment of the current state and the future potential for Per-DS; and 2) development and deployment of virtual infrastructure to support Per-DS.

5.2.2.1. *Assessment of Current State and Future Potential for Per-DS*

Given the dearth of rigorous research on Per-DS, it is incumbent on the health science and data science communities to work together to launch a comprehensive research endeavor to understand the current state and the future potential for Per-DS. One possibility would be to develop a supplemental module within the National Health and Nutrition Examination Survey (National Center for Health Statistics, 2022).

The Per-DS use survey could be structured according to the typology presented in [Section 3](#), to incorporate survey questions on respondents’ current and potential use of systematic monitoring, trigger finding, pre-post

exposure-outcome trials, and personalized trials. The ability to link the Per-DS use data to NHANES health and nutrition data will enrich the interpretation. In addition, it would be useful to include questions on the respondents' knowledge and skills in data science, as well as preference among those interested in Per-DS for study attributes such as acceptable study duration, simplicity/complexity of results presentation, and so on. While such a Per-DS use survey will need substantial planning to delineate and operationalize the key constructs, the potential payoff could be large in terms of understanding the Per-DS 'market' and how to reach its potential.

5.2.2.2. *Virtual Infrastructure to Support Per-DS*

Per-DS can realize its potential only with infrastructural support along the continuum from data collection through interpretation and application of results.

Recent advances in technology have enriched Per-DS investigators' toolbox for data acquisition. Versatile apps are available that collect objective quantitative data automatically using physiological monitors, motion sensors, geolocation, and voice analysis (Bobe, De Freitas, & Glicksberg, 2020; Daskalova et al., 2016; Estrin, 2019; Ku & Sim, 2021; Sim, 2019a). Perhaps the greater value for Per-DS studies, however, comes from convenient smartphone-based acquisition of subjective data in the form of electronic patient-reported outcomes (Sim, 2019b) and ecological momentary assessments (Shiffman et al., 2008, p. 1), that is, 'repeated sampling of subjects' current behaviors and experiences in real time, in subjects' natural environments' that capture an individual's lived experience of their health state. The focus on personal experience enables more precisely individualized data capture in Per-DS studies. Thus, smartphones, wearable devices, passive and active sensors, and their links to the expanding Internet of Things (IoT) (Ku & Sim, 2021; Sim, 2019a) present tremendous potential for Per-DS applications.

Beyond technologies whose main purpose is to *collect* data, health scientists and data scientists have collaborated to develop apps for the *design, implementation, analysis, and interpretation* of Per-DS studies. The Trialist® app used in the PREEMPT study represented a harbinger of technological developments for supporting Per-DS (Barr et al., 2015; Kravitz et al., 2018). When combined with a desktop interface and analytic backend, Trialist® allowed patients with chronic musculoskeletal pain to design their own personalized comparative effectiveness studies, track their pain-related symptoms over time, and generate results for review with the treating clinician.

In the years since Trialist® was developed, platforms for *decentralized clinical trials* have proliferated, including academic platforms such as Eureka and commercial platforms such as Evidation, Sage Bionetworks, Vibrent, Vyndiant, and CareEvolution. Decentralized clinical trials are trials that employ virtual methods for some or all aspects of clinical trials (Clinical Trials Transformation Initiative, 2021). Fully decentralized trials use web and smartphone technologies to virtually recruit, enroll, and engage thousands of participants in observational and interventional studies. Decentralized trial platforms provide modular support for structured

study design templates and data entry screens, reminders to perform data entry or to take treatments, and can be adapted to different studies, with different designs, outcomes, data structures, and presentation displays (Kravitz et al., 2020). These same components can be used for personalized studies, as shown by the I-STOP-AFib trial in which 499 participants were recruited through the Eureka platform to test individualized (N-of-1) trigger testing versus AF monitoring (Kaplan et al., 2022; Marcus et al., 2022).

Schork (2019) discussed the integration of personalized medicine, data-intensive assays, ‘big data’ research paradigms, and artificial intelligence to address the need to tailor, or ‘personalize,’ medicines to the nuanced and often unique features possessed by individual patients. Within this framework, Per-DS studies can be pursued to identify patterns in data collected on the patient that might be indicative of that patient’s response (or lack thereof) to the intervention (Schork, 2015).

Virtual trial platforms are needed to help individuals design and execute their Per-DS studies. It is certainly possible to provide scripting tools that would allow a patient with diabetes to design her own study to determine whether a lower dose of glipizide (an oral hypoglycemic) can reduce episodes of hypoglycemia without materially worsening blood sugar control. The app could randomly assign weeks for taking a higher (5 mg) versus lower (2.5 mg) dose of glipizide while integrating outcome data from a continuous glucose monitor. The app could also incorporate safety guardrails, such as not allowing more than a recommended dose, or even asking preliminary questions (how many hypoglycemic episodes have you had in past month) that trigger clinician involvement when needed. Covariates could include physical activity (input from the mobility app on the patient’s iPhone) and weight (input from her smartscale). The patient-facing output would include number and severity of hypoglycemic episodes as well as covariate-adjusted area under the glucose monitoring curve while on high- versus low-dose glipizide.

To be useful, the results returned to Per-DS investigators need to be accurate, comprehensible, and actionable while giving due attention to uncertainty. As an example of how such results can be generated, one of us (CHS) and colleagues are developing a prototype R package that fits Bayesian models and returns summaries from the posterior distribution of the treatment effect. The output from this package is fed into a mobile app to render tabular and graphical feedback to individual end users to assist their decision-making. Ultimately, this package could render its own tables and graphs and provide interpretations of results to fit into a customizable mobile app such as that described in the previous paragraph.

Providing optimal support to Per-DS investigations will necessitate using not only a variety of models suitable for different types of data (e.g., binary and continuous outcomes or correlated observations) and different questions of interest (e.g., effect modification of a treatment), but also developing flexible presentation formats that can be adapted to individual needs (see [Section 2.2](#)). For instance, results from the same analysis can be shown to the user in different ways with different interpretive language. Users comfortable with statistics could request a variety of ways to inspect and analyze the data and could even request different analyses (e.g.,

sensitivity to the choice of prior distributions). Others without this background might choose a default analysis and simple graphs with appropriate cautionary language.

Instructional videos demonstrating the capabilities of the app as well as teaching users some of the basic principles of design and analysis could increase utility and appropriate use as well as serve as scientific education for the public. As with any application that incorporates complex ideas, inappropriate use is a risk, but one that can be minimized with careful app design.

Ultimately, we need not just new apps but new ways to sort through available apps and previously run Per-DS studies to facilitate future Per-DS studies. A resource clearinghouse that indexes apps and Per-DS studies using standardized metadata and computational representation of Per-DS studies would allow any Per-DS app to run any given Per-DS study, fostering both reproducibility and replication efficiency. As an example, Jason Bobe, Joel Dudley, and colleagues at Icahn School of Medicine at Mount Sinai have developed a comprehensive virtual library of resources for personalized studies, including a lecture series, a tool to aid the design of personalized (N-of-1) studies, papers, guides, and posts (Bobe, Johnson et al., 2020). Further developments to expand the scope and reach for such virtual libraries can serve the purpose as tremendously valuable ‘virtual field guides’ to interested Per-DS investigators.

A virtual support infrastructure for Per-DS might develop into a viable business model, with a possible pathway being for health insurers (particularly managed care health plans) to develop and deploy this infrastructure as an in-house, members-only wellness toolkit for their members. Such an infrastructure might be marketable as an attractive hallmark for the insurer to recruit subscribers and members. Of course, the viability of such initiatives will depend upon the size of market demand for Per-DS investigations as well as credible evidence that Per-DS investigations lead to improved outcomes. There is fragmented evidence on the health benefits of Per-DS investigations, but more comprehensive research is needed (Cadmus-Bertram et al., 2015; Harris et al., 2019; Malhotra et al., 2018; Selker et al., 2021).

A well-established model for such a support infrastructure is the agricultural Cooperative Extension System (CES) (National Institute of Food and Agriculture, 2015). Since the CES was established under the Smith–Lever Act of 1914 (2018), it has successfully brought cutting-edge discoveries from agricultural research laboratories to farmers who can put knowledge into practice. The CES “empowers farmers, ranchers, and communities of all sizes to meet the challenges they face, adapt to changing technology, improve nutrition and food safety, prepare for and respond to emergencies, and protect our environment.”

A new Smith–Lever Act for a Per-DS Extension System might be far off, although initiatives such as the Health Advanced Research Projects Agency (HARPA) might be a possible step forward for such innovations. Smaller scale extension systems could be realized if (1) the demand for Per-DS is indeed large and sustainable and (2) evidence for the health benefits of Per-DS investigations continue to emerge and coalesce.

6. Summary and Discussion

The history of scientific discovery since the Enlightenment has privileged the position of ‘experts’ who had the time, leisure, and resources not only to make discoveries, but also to determine the questions to be asked. Personalized data science (Per-DS) upends this paradigm by putting the individual at the center. As we have shown, implementing this approach requires new ways of thinking—expanding the role of individuals not only as participants, but also as investigators and decision makers; and envisioning new roles for experts not as controllers, but as guidance counselors. However, in many ways, Per-DS reflects a rich human tradition of curiosity, self-discovery, and self-experimentation, tracking back to the earliest hunter-gatherers who must have performed Per-DS investigations to discover the nutritional value, appeal, and toxicity of various types of food. One example is the legend of Emperor Shennong’s investigations into the medicinal effects of a large number of herbs, discussed further in [Appendix C, “Personalized Data Science \(Per-DS\) through Millennia: Shennong, Barry Marshall, Daniel Carrión, and Beyond.”](#)

6.1. Barriers and Facilitators for Per-DS Studies

Despite its promising potential, the realization of Per-DS studies needs to overcome numerous barriers, including the lack of infrastructure, the inability to access on-device data, the need for financing/insurance coverage, and the burden of human subjects protection regulations. Important facilitators are also needed, including public health education to inform prospective Per-DS investigators, and medical education to inform clinicians.

Lack of infrastructure is the most crucial barrier for Per-DS studies. While many individuals might be interested in conducting Per-DS studies to inform their own health and wellness decisions, few of them have access to the infrastructure needed to implement their own studies safely and effectively. In particular, Per-DS investigators need support from peers, clinicians, the data science and subject-area science communities, and from both government and business sources. Particular needs include infrastructure support to develop and deploy apps that facilitate Per-DS investigations; a resource clearinghouse for Per-DS investigators to search comprehensively for information and material; and tools/apps to enable the conduct of their Per-DS investigations. Such a virtual infrastructure might be modeled after the very successful agricultural Cooperative Extension System (National Institute of Food and Agriculture, 2015).

A barrier closely related to the lack of infrastructure is the inability of patients to access their on-device data, such as data recorded on their PAP devices or CGMs. While patients might think they own the data on their medical devices, device manufacturers might restrict their data access using technological protection measures (TPMs), such as encryption. Patient data access therefore involves complex copyright issues governed by Digital Millennium Copyright Act of 1998 (DMCA) (U.S. Copyright Office, 1998). Section 1201 of DMCA forbids circumvention of TPMs, thus blocking patient access to protected data on their medical devices, while also delegating authority to the Library of Congress and the U.S. Copyright Office to conduct triennial reviews

and promulgate exemptions for classes of work that necessitate such exemptions. Each exemption is valid for 3 years, and can be reapplied and renewed if approved.

In 2015, the Library of Congress promulgated a 3-year exemption to Section 1201, allowing patients to access data generated by their medical devices that were wholly or partially implanted in their bodies (Sellars, 2015; U.S. Copyright Office 2015). This exemption was renewed for 3 more years in 2018 (U.S. Copyright Office, 2018). In 2021, this exemption was renewed and expanded to remove the restriction to implanted medical devices, and also expanded to allow patient's agents to access their data, thereby allowing patient access to their on-device data for the next 3 years (U.S. Copyright Office, 2021).

Financing the creation and operation of infrastructure to support Per-DS investigations will not be cheap or easy. But then again, limiting individuals' ability to generate personally useful data has its own costs. In clinical settings, evidence derived from Pop-DS will not always match the characteristics, needs, and values of individual patients. The usual solution—'informal trials of therapy'—can lead to inaccurate results. In community settings, people are making all sorts of personal health decisions without the benefit of rigorous and scientifically valid empirical information. Therefore, investments in infrastructure support for Per-DS might yield substantial returns that justify the investments.

Insurance coverage for Per-DS investigations will help facilitate the conduct of Per-DS studies. Existing insurance policies might need to be modified to incentivize the conduct of Per-DS studies to inform personal treatment or health decisions. Virtual support infrastructure for Per-DS could potentially develop into a viable product for health insurers to deploy as an attractive hallmark to recruit subscribers and members.

The burden of human subjects protection regulations can be prohibitive to the conduct of Per-DS studies, given that they resemble research studies. It is therefore important to recognize that a Per-DS study that aims to produce personal knowledge to serve the individual's personal needs does not qualify for the U.S. Department of Health and Human Services definition for research, and thus should be exempt from human subjects protections regulations for research studies that aim to produce generalizable knowledge ([Section 4.4.2](#)).

Once infrastructure is available and regulatory barriers are removed, developing a robust market for Per-DS studies will require a combination of support, education, and persuasion. For individuals interested in such studies but lacking knowledge and skill to proceed, a 'pull' strategy emphasizing support ('how to' videos, telephone helplines, etc.) may be adequate, as these individuals are ready and willing consumers who simply need activating.

For individuals who are more reticent, a proactive 'push' strategy might be needed. Such efforts should incorporate both education and persuasion. The COVID-19 pandemic has highlighted the need to enhance science literacy in the United States and elsewhere (Braund, 2021; World Health Organization, 2020). Deficits are pervasive, implying a need for much greater investment in science education, starting with the early grades.

Ironically, Per-DS studies could themselves be part of the solution, as engaging schoolchildren in the process of data collection and analysis on topics that affect their own lives could be a powerful educational experience. However, a comprehensive ‘push’ strategy will also need to incorporate persuasion, channeled through direct-to-consumer advertising and social networks.

Given the uneven distribution of science literacy, such a ‘push’ effort is likely to engage some members of the population but not others, particularly those with strong mistrust of science overall and of the evidence-based decision paradigm in particular. We need to recognize that Per-DS, like almost all public health programs, should not expect to achieve 100% uptake.

In the long term, public health education efforts incorporating Per-DS studies in the early grades, as suggested above, might be one way to bring even skeptical members of the public onboard—essentially showing them how science can benefit them as individuals via protocols under *their* control and not dictated by scientific or medical elites or by the government. Whether efforts like these will succeed in engaging a large segment of the population is an empirical question worthy of further study (as we argue in [Section 5.2.2.1](#)).

In the foreseeable future, health and data professionals might need to navigate the opportunities and challenges of Per-DS, to work collaboratively with patients interested in Per-DS investigations, and to refer patients to quality health apps for assistance in performing Per-DS functions. Because physicians will be an important component of the Per-DS ‘infrastructure,’ these topics deserve to be incorporated into the medical school curriculum. A simple but potentially effective active-learning approach would be to require that every medical student design, participate in, and analyze/interpret at least one personalized trial that is relevant to their own personal health. From such beginnings, a robust field could emerge, bearing fruit for individual well-being, scientific progress, and more inclusive and less hierarchical decision making.

6.2. Hybrid Models Beyond ‘Pure’ Per-DS

This article emphasizes a ‘pure’ model for Per-DS, focused on investigations conducted by an individual (maybe with help from professionals in health/wellness and/or data science) using her own data to inform her health decisions, instead of in a group of individuals sharing data. We now provide a cursory discussion of hybrid models that consider the individual Per-DS investigator in a broader context.

Individual Per-DS investigators will likely benefit from peers who have conducted or are in the process of conducting their own Per-DS investigations. Like home gardeners, peer support can help the novice Per-DS investigator to learn from the experience of others regarding what works, what does not work, what pitfalls need to be watched out for, and so on. Such peer interactions can be facilitated with a virtual peer forum that should be part of the virtual infrastructure discussed in [Section 5.2.2.2](#), potentially supported by the clinical practice or the health care system as part of their innovative business model to enhance efficiency, impact, and cost-effectiveness.

It is conceivable such a peer network can be developed further to share not only ‘tips’ about conducting the Per-DS investigations, but also to help Per-DS investigators share data across similar Per-DS studies. Such data sharing can yield benefits for each individual Per-DS investigator, by ‘borrowing from strength’ to improve the precision for the results, while balancing the bias–variance trade-off. It is also conceivable that such data sharing could result in generalizable knowledge that can benefit others, making it attractive to researchers, clinical practices, and health care systems supporting the Per-DS endeavor.

The potential benefits of data sharing would need to be balanced carefully against any potential loss of individual-centered autonomy manifested in the ‘pure’ Per-DS model. For example,

- While the ‘pure’ Per-DS model emphasizes personalization of the protocol for each individual Per-DS study, the hybrid model might standardize the protocol to render comparable the data produced from individual trials, to allow data to be pooled across individuals for statistical analyses such as Bayesian meta-analysis.
- While the ‘pure’ Per-DS model emphasizes the production of personal knowledge for use by each individual Per-DS investigator (and should therefore be exempt from human subjects regulations), the hybrid model might be contingent upon human subjects regulations, especially if the data are pooled across individuals to produce generalizable knowledge to benefit others. For instance, health care organizations might wish to educate clinicians by having them enroll their patients in individual trials that could then inform practice guidance. To the extent that the sponsoring organizations want to collect and analyze data from Per-DS studies to inform practice, the hybrid model might be deemed as research.

Therefore, there is a range of studies that span the continuum between the ‘pure’ Per-DS model and the Pop-DS model, such as:

- A consortium of Per-DS investigations might retain the key features of the Per-DS model, including the focus on each individual as the investigator, study participant, and beneficiary, allowing individual Per-DS investigators to exercise preference over the design and implementation of their own studies, while facilitating data sharing to borrow from strength to benefit individual investigators, and also to produce generalizable knowledge to benefit others who did not participate in the consortium. Such an endeavor should be recognized as a hybrid model for Per-DS with partial allowance for Pop-DS through the pooling of the data across individuals.
- A cohort study might use methods such as personalized trials, but designed and managed centrally with limited or no personalization for the design and implementation of each participant’s protocol, aiming primarily to produce generalizable knowledge to benefit others who did not participate in the study. While such a study is more like a Pop-DS study than a Per-DS study, it might still yield valuable benefits for the individual participants, if the study analyzes each participant’s data (possibly borrowing from strength with other participants’ data) and shares individual-specific findings (personal knowledge) with each individual participant. Such an endeavor should be recognized as a hybrid model for Pop-DS with partial allowance for Per-DS through the provision of personal knowledge to individual participants.

The methodology for the hybrid model for Per-DS entails a complicated exercise in meta-analysis for the data combined across individuals, as it needs to accommodate the personalized design and implementation of individual Per-DS investigations ([Section 2](#)). Further development and deployment of this hybrid methodology using meta-regression (Schmid, 1999; Schmid et al., 2004; Wallace et al., 2009) and network meta-analysis (Efthimiou et al., 2016; Hutton et al., 2015; Kaplan et al., 2022; Marcus et al., 2022; Schmid et al., 2014; Wang, 2020) is an important expansion of the work reported in this article. The article by Schmid and Yang (2022) in this issue provides important guidance for this approach.

6.3. Spillover Benefits Within and Beyond Health/Wellness

Aside from the potential benefits to individuals, Per-DS could have spillover benefits to enhance scientific literacy in the general population by engaging a broad, grassroots constituency to take an active role to conduct their own scientific investigations. Hands-on experience with study design and scientific methods could help people develop an empirical mindset with a deeper appreciation for science in general, more skepticism about claims for products and policies that are not supported by rigorous science, and perhaps an understanding of how not to be manipulated by misinformation (Wheeler, 1976). Some Per-DS investigators might be empowered to take a proactive role in pursuing credible evidence beyond their own specific Per-DS investigations. They might even transfer their proactive role to other domains in life. The impact of those spillover benefits (‘externalities’) might be powerful and long lasting.

Acknowledgments

We appreciate Hugo Campos, Ying Kuen (Ken) Cheung, Karina Davidson, Mark Drangsholt, Charles Duan, Deborah Estrin, ChihMing Fan, Xiao-Li Meng, Jef Pearlman, Mary Jane Rotheram, Jürgen Unützer, and Sunita Vohra for fruitful discussions that helped shape this article; ChihMing Fan and Laura Hyppolite for assistance with the preview image, and constructive comments from the editor and two reviewers who helped improve the thoroughness and clarity of this article.

Disclosure Statement

This work was supported by grants R01LM012836 from the National Library of Medicine of the National Institutes of Health and P30AG063786 from the National Institute on Aging of the National Institutes of Health. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication. The views expressed in this paper are those of the authors and do not represent the views of the National Institutes of Health, the U.S. Department of Health and Human Services, or any other government entity.

Dr. Richard Kravitz is an advisor to and holds stock options from Sam Dot Company, a firm involved in developing information technology infrastructure for personalized trials. Dr. Ida Sim was a co-investigator for Eureka, and is a co-investigator with Sage Bionetworks. The other authors have no disclosures to share for this manuscript.

Mention of product names or brand names is for information only and does not indicate endorsement.

References

American Diabetes Association. (2020, December 23). *CGM & time in range*. <https://www.diabetes.org/tools-support/devices-technology/cgm-time-in-range>

Apnea Board. (2019, April 23). *OSCAR – Open Source CPAP Analysis Reporter*. Retrieved March 13, 2022, from <http://www.apneaboard.com/forums/Thread-OSCAR-Open-Source-CPAP-Analysis-Reporter--25239>

Apnea Board. (2011, March 5). *Apnea Board: Sleep apnea patients helping one another*. <https://www.apneaboard.com/>

Austin, C. P., Mount, B. A., & Colvis, C. M. (2021). Envisioning an actionable research agenda to facilitate repurposing of off-patent drugs. *Nature Reviews Drug Discovery*, 20(10), 723–724. <https://doi.org/10.1038/d41573-021-00090-y>

Barr, C., Marois, M., Sim, I., Schmid, C. H., Wilsey, B., Ward, D., Duan, N., Hays, R. D., Selsky, J., Servadio, J., Schwartz, M., Dsouza, C., Dhammi, N., Holt, Z., Baquero, V., MacDonald, S., Jerant, A., Sprinkle, R., & Kravitz, R. L. (2015). The PREEMPT study – Evaluating smartphone-assisted n-of-1 trials in patients with chronic pain: study protocol for a randomized controlled trial. *Trials*, 16, Article 67. <https://doi.org/10.1186/s13063-015-0590-8>

Bobe, J. R., De Freitas, J. K., & Glicksberg, B. S. (2020). Exploring the potential for collaborative use of an app-based platform for N-of-1 trials among healthcare professionals that treat patients with insomnia. *Frontiers in Psychiatry*, 11, Article 530995. <https://doi.org/10.3389/fpsy.2020.530995>

Bobe, J., Johnson, M., Jones, M., & Viglizzo, R. (2020, December 2). *N-of-1 resources: Guides, papers, and posts on N-of-1*. <https://www.n1app.org/n1-resources>

Bonner, C., Trevena, L. J., Gaissmaier, W., Han, P., Okan, Y., Ozanne, E., Peters, E., Timmermans, D., & Zikmund-Fisher, B. J. (2021). Current best practice for presenting probabilities in patient decision aids: Fundamental principles. *Medical Decision Making*, 41(7), 821–833. <https://doi.org/10.1177/0272989X21996328>

- Braund, M. (2021). Critical STEM literacy and the COVID-19 pandemic. *Canadian Journal of Science, Mathematics and Technology Education*, 21(2), 339–356. <https://doi.org/10.1007/s42330-021-00150-w>
- Cadmus-Bertram, L. A., Marcus, B. H., Patterson, R. E., Parker, B. A., & Morey, B. L. (2015). Randomized trial of a fitbit-based physical activity intervention for women. *American Journal of Preventive Medicine*, 49(3), 414–418. <https://doi.org/10.1016/j.amepre.2015.01.020>
- Center for Disease Control and Prevention. (2022). *Symptoms of COVID-19: Watch for symptoms*. Retrieved July 4, 2022, from <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>
- Chen, C., Haddad, D., Selsky, J., Hoffman, J. E., Kravitz, R. L., Estrin, D. E., & Sim, I. (2012). Making sense of mobile health data: An open architecture to improve individual- and population-level health. *Journal of Medical Internet Research*, 14(4), Article e112. <https://doi.org/10.2196/jmir.2152>
- Cheung, K., & Mitsumoto, H. (2022). Evaluating personalized (N-of-1) trials in rare diseases: How much experimentation is enough? *Harvard Data Science Review*, (Special Issue 3). <https://doi.org/10.1162/99608f92.e11adff0>
- Clinical Trials Transformation Initiative. (2021). Digital health trials: Recommendations for supporting decentralized trial approaches. Retrieved March 13, 2022, from https://ctti-clinicaltrials.org/wp-content/uploads/2021/07/CTTI_Supporting_Decentralized_Approaches_Recommendations.pdf
- Columbia University Data Science Institute. (2020, September 22). *About DSI: Data science is the study of extracting value from data*. Retrieved January 3, 2022, from <https://datascience.columbia.edu/about-us/>
- Dahlhamer, J., Lucas, J., Zelaya, C., Nahin, R., Mackey, S., DeBar, L., Kerns, R., Von Korff, M., Porter, L., & Helmick, C. (2018). Prevalence of chronic pain and high-impact chronic pain among adults—United States, 2016. *MMWR Morbidity and Mortality Weekly Report*, 67(36), 1001–1006. <https://doi.org/10.15585/mmwr.mm6736a2>
- Daskalova, N., Metaxa-Kakavouli, D., Tran, A., Nugent, N., Boergers, J., McGeary, J., & Huang, J. (2016). SleepCoacher: A personalized automated self-experimentation system for sleep recommendations. In J. Rekimoto & Takeo Igarashi (Eds.), *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)* (pp. 347–358). Association for Computing Machinery. <https://doi.org/10.1145/2984511.2984534>
- David, L. A., Materna, A. C., Friedman, J., Campos-Baptista, M. I., Blackburn, M. C., Perrotta, A., Erdman, S. E., & Alm, E. J. (2014). Host lifestyle affects human microbiota on daily timescales. *Genome Biology*, 15(7), Article R89. <https://doi.org/10.1186/gb-2014-15-7-r89>

Davidson, K. W., Silverstein, M., Cheung, K., Paluch, R. A., & Epstein, L. H. (2021). Experimental designs to optimize treatments for individuals: Personalized N-of-1 trials. *JAMA Pediatrics*, 175(4), 404–409. <https://doi.org/10.1001/jamapediatrics.2020.5801>.

Derby, L., Kronish, I. M., Wood, D., Cheung, Y., Cohn, E., Duan, N., St Onge, T., Duer-Hefe, J., Davidson, K. W., & Moise, N. (2021). Using a multistakeholder collaboratory and patient surveys to inform the conduct of personalized (N-of-1) trials. *Health Psychology*, 40(4), 230–241. <https://doi.org/10.1037/hea0001058>

Dominici, F., & Parkes, D. C. (2021). Introducing the interim co-editors-in-chief. *Harvard Data Science Review*, 3(3). <https://doi.org/10.1162/99608f92.24ee17a7>

Drangsholt, M. (2016, January 24). Learning from my N of 1. *Quantified Self Public Health*. <https://medium.com/quantified-self-public-health/learning-from-my-n-of-1-6d1dcab32d4a>

Duan, N., Kravitz, R. L., & Schmid, C. H. (2013). Single-patient (n-of-1) trials: A pragmatic clinical decision methodology for patient-centered comparative effectiveness research. *Journal of Clinical Epidemiology*, 66(8 Suppl), S21–S28. <https://doi.org/10.1016/j.jclinepi.2013.04.006>

Efthimiou, O., Debray, T. P., van Valkenhoef, G., Trelle, S., Panayidou, K., Moons, K. G., Reitsma, J. B., Shang, A., Salanti, G., & GetReal Methods Review Group (2016). GetReal in network meta-analysis: A review of the methodology. *Research Synthesis Methods*, 7(3), 236–263. <https://doi.org/10.1002/jrsm.1195>

Estrin, D. (2013). What happens when each patient becomes their own "universe" of unique medical data? [Talk]. TEDMED. <https://www.tedmed.com/talks/show?id=17762>

Estrin, D. (2014). Small data, where N = me. *Communications of the ACM*, 57(4), 32–34. <https://doi.org/10.1145/2580944>

Estrin, D. (2019, February 4–5). *Small data for better health: Technologies for personalizing assessments and interventions* [Keynote address]. Scientific Meeting on Big Data for Better Science, Royal Society of London.

Forsdyke D. R. (2015). Summertime dosage-dependent hypersensitivity to an angiotensin II receptor blocker. *BMC Research Notes*, 8, Article 227. <https://doi.org/10.1186/s13104-015-1215-8>

Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The science of visual data communication: What works. *Psychological Science in the Public Interest*, 22(3), 110–161. <https://doi.org/10.1177/15291006211051956>

Harris, T., Limb, E. S., Hosking, F., Carey, I., DeWilde, S., Furness, C., Wahlich, C., Ahmad, S., Kerry, S., Whincup, P., Victor, C., Ussher, M., Iliffe, S., Ekelund, U., Fox-Rushby, J., Ibison, J., & Cook, D. G. (2019). Effect of pedometer-based walking interventions on long-term health outcomes: Prospective 4-year follow-up

of two randomised controlled trials using routine primary care data. *PLoS Medicine*, 16(6), Article e1002836. <https://doi.org/10.1371/journal.pmed.1002836>

Health eHeart Study. (2012, June 11). *Join the study to end heart disease*. <https://www.health-eheartstudy.org/>

Henson, P., David, G., Albright, K., & Torous, J. (2019). Deriving a practical framework for the evaluation of health apps. *The Lancet Digital Health*, 1(2), e52–e54. [https://doi.org/10.1016/S2589-7500\(19\)30013-5](https://doi.org/10.1016/S2589-7500(19)30013-5)

Howden, L. M., & Meyer, J. A. (2011, May 1). *Age and Sex Composition: 2010*. Report no. C2010BR-03. U.S. Census Bureau. <https://www.census.gov/library/publications/2011/dec/c2010br-03.html>

Hsieh, C. K., Alquaddoomi, F., Okeke, F., Pollak, J. P., Gunasekara, L., & Estrin, D. (2018). Small data: Applications and architecture. In G. Weckman & J. Grzymala-Busse (Eds.), *Proceedings of ALLDATA 2018: The Fourth International Conference on Big Data, Small Data, Linked Data and Open Data* (pp. 1–10). IARIA. https://destrin.tech.cornell.edu/papers/small_data_applications_alldata_2018.pdf

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., Ioannidis, J. P., Straus, S., Thorlund, K., Jansen, J. P., Mulrow, C., Catalá-López, F., Gøtzsche, P. C., Dickersin, K., Boutron, I., Altman, D. G., & Moher, D. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Annals of Internal Medicine*, 162(11), 777–784. <https://doi.org/10.7326/M14-2385>

ImproveCareNow. (2020, November 9). ICN Registry. https://www.improvecarenow.org/icn_registry_digest

Kale, A., Kay, M., & Hullman, J. (2021). Visual reasoning strategies for effect size judgments and decisions. *IEEE Transactions on Visualization & Computer Graphics*, 27(2), 272–282. <https://doi.org/10.1109/TVCG.2020.3030335>

Kaplan, H., Marcus, G. M., Schmid, C. H., Schuler, C. L., Dodds, C. M., Murphy, L. E., Modrow, M. F., Sigona, K., & Opiari-Arrigan, L. (2022). *Using single subject (N-of-1) designs to answer patient-identified research questions*. Report no. PCORI-D-21-00076. Patient-Centered Outcomes Research Institute.

Kelly, S. (2021, October 13). Abbott, Dexcom could benefit from CGM use jump among Type 2 diabetes patients: Report. *MedTech Dive*. <https://www.medtechdive.com/news/abbott-dexcom-diabetes-tech-type-2-jefferies/608118/>

Kent, D. M., & Hayward, R. A. (2007). Limitations of applying summary results of clinical trials to individual patients: The need for risk stratification. *JAMA*, 298(10), 1209–1212. <https://doi.org/10.1001/jama.298.10.1209>

Kent, D. M., Paulus, J. K., van Klaveren, D., D'Agostino, R., Goodman, S., Hayward, R., Ioannidis, J., Patrick-Lake, B., Morton, S., Pencina, M., Raman, G., Ross, J. S., Selker, H. P., Varadhan, R., Vickers, A., Wong, J. B.,

& Steyerberg, E. W. (2020). The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement. *Annals of Internal Medicine*, 172(1), 35–45. <https://doi.org/10.7326/M18-3667>

Kent, D. M., Steyerberg, E., & van Klaveren, D. (2018). Personalized evidence based medicine: Predictive approaches to heterogeneous treatment effects. *BMJ*, 363, Article k4245. <https://doi.org/10.1136/bmj.k4245>

Kent, D. M., van Klaveren, D., Paulus, J. K., D'Agostino, R., Goodman, S., Hayward, R., Ioannidis, J., Patrick-Lake, B., Morton, S., Pencina, M., Raman, G., Ross, J. S., Selker, H. P., Varadhan, R., Vickers, A., Wong, J. B., & Steyerberg, E. W. (2020). The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement: Explanation and Elaboration. *Annals of Internal Medicine*, 172(1), W1–W25. <https://doi.org/10.7326/M18-3668>

Kim, Y. S., Kayongo, P., Grunde-McLaughlin, M., & Hullman, J. (2021) Bayesian-assisted inference from visualized data. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 989–999. <https://doi.org/10.1109/TVCG.2020.3028984>

Kravitz, R. (1990). Serving several masters: Conflicting responsibilities in health services research. *Journal of General Internal Medicine*, 5(2), 170–174. <https://doi.org/10.1007/BF02600521>

Kravitz, R. L., Aguilera, A., Chen, E. J., Choi, Y. K., Hekler, E., Karr, C., Kim, K. K., Phatak, S., Sarkar, S., Schueller, S. M., Sim, I., Yang, J., & Schmid, C. H. (2020). Feasibility, acceptability, and influence of mhealth-supported N-of-1 trials for enhanced cognitive and emotional well-being in US volunteers. *Frontiers in Public Health*, 8, Article 260. <https://doi.org/10.3389/fpubh.2020.00260>

Kravitz, R. L., & Duan, N. (2022). Conduct and implementation of personalized trials in research and practice. *Harvard Data Science Review*, (Special Issue 3). <https://doi.org/10.1162/99608f92.901255e7>

Kravitz, R., Duan, N., & Braslow, J. T. (2004). Evidence based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Quarterly*, 82(4), 661–687. Erratum in *Milbank Quarterly* 2006, 84(4), 759–760. <https://doi.org/10.1111/j.0887-378x.2004.00327.x>

Kravitz, R. L., & Duan, N., eds., and the DEcIDE Methods Center N-of-1 Guidance Panel (Duan, N., Eslick, I., Gabler, N. B., Kaplan, H. C., Kravitz, R. L., Larson, E. B., Pace, W. D., Schmid, C. H., Sim, I., & Vohra, S.). (2014). *Design and implementation of N-of-1 trials: A user's guide*. Publication No. 13(14)-EHC122-EF. Agency for Healthcare Research and Quality. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/n-1-trials_research-2014-5.pdf.

Kravitz, R. L., Paterniti, D. A., Hay, M. C., Subramanian, S., Dean, D. E., Weisner, T., Vohra, S., & Duan, N. (2009). Marketing therapeutic precision: Potential facilitators and barriers to adoption of n-of-1 trials. *Contemporary Clinical Trials*, 30(5), 436–445. <https://doi.org/10.1016/j.cct.2009.04.001>

- Kravitz, R. L., Schmid, C. H., Marois, M., Wilsey, B., Ward, D., Hays, R. D., Duan, N., Wang, Y., MacDonald, S., Jerant, A., Servadio, J. L., Haddad, D., & Sim, I. (2018). Effect of mobile device-supported single-patient multi-crossover trials on treatment of chronic musculoskeletal pain: A randomized clinical trial. *JAMA Internal Medicine*, 178(10), 1368–1377. <https://doi.org/10.1001/jamainternmed.2018.3981>
- Ku, J. P., & Sim, I. (2021). Mobile health: Making the leap to research and clinics. *NPJ Digital Medicine*, 4(1), Article 83. <https://doi.org/10.1038/s41746-021-00454-z>
- Lagan, S., Aquino, P., Emerson, M. R., Fortuna, K., Walker, R., & Torous, J. (2020). Actionable health app evaluation: Translating expert frameworks into objective metrics. *NPJ Digital Medicine*, 3(1), Article 100. <https://doi.org/10.1038/s41746-020-00312-4>
- Lee, A. W., Kenfield, S. A., Wang, E. Y., Enriquez, A., Oni-Orisan, A., Steinman, M. A., Sim, I., Breyer, B. N., & Bauer, S. R. (2021). Tracking lower urinary tract symptoms and tamsulosin side effects among older men using a mobile app (PERSONAL): Feasibility and usability study. *JMIR Formative Research*, 5(12), Article e30762. <https://doi.org/10.2196/30762>
- Li, X., Dunn, J., Salins, D., Zhou, G., Zhou, W., Schüssler-Fiorenza Rose, S. M., Perelman, D., Colbert, E., Runge, R., Rego, S., Sonecha, R., Datta, S., McLaughlin, T., & Snyder, M. P. (2017). Digital health: Tracking physiomes and activity using wearable biosensors reveals useful health-related information. *PLoS Biology*, 15(1), Article e2001402. <https://doi.org/10.1371/journal.pbio.2001402>
- Li, X., & Meng, X.-L. (2021) A multi-resolution theory for approximating infinite- p -Zero- n : Transitional inference, individualized predictions, and a world without bias-variance tradeoff, *Journal of the American Statistical Association*, 116(533), 353–367. <https://doi.org/10.1080/01621459.2020.1844210>
- Lindstrom, M. (2016). *Small data: The tiny clues that uncover huge trends*. St. Martin's Press.
- Liu, K., & Meng, X. L. (2016). There is individualized treatment. Why not individualized inference? *Annual Review of Statistics and Its Application*, 3, 79–111. <https://doi.org/10.1146/annurev-statistics-010814-020310>
- Maclure, M. (1991). The case-crossover design: A method for studying transient effects on the risk of acute events. *American Journal of Epidemiology*, 133(2), 144–153. <https://doi.org/10.1093/oxfordjournals.aje.a115853>
- Maclure, M., & Mittleman, M. A. (2000). Should we use a case-crossover design? *Annual Review of Public Health*, 21, 193–221. <https://doi.org/10.1146/annurev.publhealth.21.1.193>
- Malhotra, A., Crocker, M. E., Willes, L., Kelly, C., Lynch, S., & Benjafield, A. V. (2018). Patient engagement using new technology to improve adherence to positive airway pressure therapy: A retrospective analysis. *Chest*, 153(4), 843–850. <https://doi.org/10.1016/j.chest.2017.11.005> .

Marcus, G. M., Modrow, M. F., Schmid, C. H., Sigona, K., Nah, G., Yang, J., Chu, T. C., Joyce, S., Gettabecha, S., Ogomori, K., Yang, V., Butcher, X., Hills, M. T., McCall, D., Sciarappa, K., Sim, I., Pletcher, M. J., & Olgin, J. E. (2022). Individualized studies of triggers of paroxysmal atrial fibrillation: The I-STOP-AFib randomized clinical trial. *JAMA Cardiology*, 7(2), 167–174. <https://doi.org/10.1001/jamacardio.2021.5010>

Meng, X.-L. (2014). A trio of inference problems that could win you a Nobel Prize in statistics (if you help fund it). In X. Lin, C. Genest, D. L. Banks, G. Molenberghs, D. W. Scott, & J.-L. Wang (Eds.), *Past, present, and future of statistical science* (pp. 537–562). CRC Press.

Meng, X.-L. (2018) Statistical paradises and paradoxes in big data (I): Law of large populations, big data paradox, and the 2016 US presidential election. *Annals of Applied Statistics*, 12(2), 685–726. <https://doi.org/10.1214/18-AOAS1161SF>

Mohr, D. C., Cheung, K., Schueller, S. M., Hendricks Brown, C., & Duan, N. (2013). Continuous evaluation of evolving behavioral intervention technologies. *American Journal of Preventive Medicine*, 45(4), 517–523. <https://doi.org/10.1016/j.amepre.2013.06.006>

Mohr, D. C., Schueller, S. M., Riley, W. T., Brown, C. H., Cuijpers, P., Duan, N., Kwasny, M. J., Stiles-Shields, C., & Cheung, K. (2015). Trials of intervention principles: Evaluation methods for evolving behavioral intervention technologies. *Journal of Medical Internet Research*, 17(7), Article e166. <https://doi.org/10.2196/jmir.4391>

Moise, N., Wood, D., Cheung, Y., Duan, N., Onge, T. S., Duer-Hefelee, J., Pu, T., Davidson, K. W., Kronish, I. M., & Members of the “Personalized Trial Collaboratory.” (2018). Patient preferences for personalized (N-of-1) trials: A conjoint analysis. *Journal of Clinical Epidemiology*, 102, 12–22. <https://doi.org/10.1016/j.jclinepi.2018.05.020>

National Center for Health Statistics. (2022). *National Health and Nutrition Examination Survey*. Retrieved March 13, 2022, from <https://www.cdc.gov/nchs/nhanes/index.htm>

National Institute of Food and Agriculture. (2015, February 15). *Cooperative Extension System*. <https://nifa.usda.gov/cooperative-extension-system>

Peppard, P. E., Young, T., Barnet, J. H., Palta, M., Hagen, E. W., & Hla, K. M. (2013). Increased prevalence of sleep-disordered breathing in adults. *American Journal of Epidemiology*, 177(9), 1006–1014. <https://doi.org/10.1093/aje/kws342>

Phillips, C. J. (2020). Precision medicine and its imprecise history. *Harvard Data Science Review*, 2(1). <https://doi.org/10.1162/99608f92.3e85b56a>

Protection of Human Subjects, 45 C.F.R. Part 46 (2022). <https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46>

Punja, S., Eslick, I., Duan, N., Vohra, S., & the DEcIDE Methods Center N-of-1 Guidance Panel. (2014). An ethical framework for N-of-1 trials: Clinical care, quality improvement, or human subjects research? In R. L. Kravitz & N. Duan (Eds.), and the DEcIDE Methods Center N-of-1 Guidance Panel, *Design and implementation of N-of-1 trials: A user's guide* (pp. 13–22). Publication No. 13(14)-EHC122-EF. Agency for Healthcare Research and Quality. <https://effectivehealthcare.ahrq.gov/products/n-1-trials/research-2014-3>

Quantified Self. (n.d.). *What is Quantified Self?* Retrieved March 13, 2022, from <https://quantifiedself.com/about/what-is-quantified-self/>

Saint-Maurice, P. F., Troiano, R. P., Bassett, D. R., Jr, Graubard, B. I., Carlson, S. A., Shiroma, E. J., Fulton, J. E., & Matthews, C. E. (2020). Association of daily step count and step intensity with mortality among US adults. *JAMA*, 323(12), 1151–1160. <https://doi.org/10.1001/jama.2020.1382>

Schandelmaier, S., Briel, M., Varadhan, R., Schmid, C. H., Devasenapathy, N., Hayward, R. A., Gagnier, J., Borenstein, M., van der Heijden, G., Dahabreh, I. J., Sun, X., Sauerbrei, W., Walsh, M., Ioannidis, J., Thabane, L., & Guyatt, G. H. (2020). Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *CMAJ*, 192(32), E901–E906. <https://doi.org/10.1503/cmaj.200077>

Schmid, C. H. (1999). Exploring heterogeneity in randomized trials via meta-analysis. *Therapeutic Innovation & Regulatory Science*, 33, 211–224. <https://doi.org/10.1177/009286159903300124>

Schmid C. H., Landa, M., Jafar, T. H., Giatras, I., Karim, T., Reddy, M., Stark, P. C., & Levey, A. S. for the Angiotensin-Converting Enzyme Inhibition in Progressive Renal Disease (AIPRD) Study Group. (2003). Constructing a database of individual clinical trials for longitudinal analysis. *Controlled Clinical Trials*, 24(3), 324–340. [https://doi.org/10.1016/s0197-2456\(02\)00319-7](https://doi.org/10.1016/s0197-2456(02)00319-7)

Schmid, C. H., Stark, P. C., Berlin, J. A., Landais, P., & Lau, J. (2004). Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors. *Journal of Clinical Epidemiology*, 57(7), 683–697. <https://doi.org/10.1016/j.jclinepi.2003.12.001>

Schmid, C. H., Trikalinos, T. A., & Olkin, I. (2014). Bayesian network meta-analysis for unordered categorical outcomes with incomplete data. *Research Synthesis Methods*, 5(2), 162–185. <https://doi.org/10.1002/jrsm.1103>

Schmid, C. H., & Yang, J. (2022). Bayesian models for N-of-1 trials. *Harvard Data Science Review*, (Special Issue 3). <https://doi.org/10.1162/99608f92.3f1772ce>

- Schork, N. J. (2015). Personalized medicine: Time for one-person trials. *Nature*, 520(7549), 609–611. <https://doi.org/10.1038/520609a>
- Schork, N. J. (2019). Artificial intelligence and personalized medicine. *Cancer Treatment and Research*, 178, 265–283. https://doi.org/10.1007/978-3-030-16391-4_11
- Schwartz, B., & Ward, A. (2004). Doing better but feeling worse: The paradox of choice. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 86–104). John Wiley & Sons.
- Schwartz, D., & Lellouch, J. (1967). Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Diseases*, 20(8), 637–648. [https://doi.org/10.1016/0021-9681\(67\)90041-0](https://doi.org/10.1016/0021-9681(67)90041-0)
- Schwartz, D., & Lellouch, J. (2009). Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Clinical Epidemiology*, 62(5), 499–505. Reprint of Schwartz & Lellouch (1967). <https://doi.org/10.1016/j.jclinepi.2009.01.012>
- Selker, H. P., Cohen, T., D'Agostino, R. B., Dere, W. H., Ghaemi, S. N., Honig, P. K., Kaitin, K. I., Kaplan, H. C., Kravitz, R. L., Larholt, K., McElwee, N. E., Oye, K. A., Palm, M. E., Perfetto, E., Ramanathan, C., Schmid, C. H., Seyfert-Margolis, V., Trusheim, M., & Eichler, H. G. (2021). A useful and sustainable role for N-of-1 trials in the healthcare ecosystem. *Clinical Pharmacology and Therapeutics*. Advance online publication. <https://doi.org/10.1002/cpt.2425>
- Sellars, A. (2015). *DMCA exemption granted for med device research, patient access to data*. <https://clinic.cyber.harvard.edu/2015/10/27/dmca-exception-granted-for-medical-device-research-patient-access-to-data/>
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 1–32. <https://doi.org/10.1146/annurev.clinpsy.3.022806.091415>
- Sim, I. (2019a). Mobile devices and health. *The New England Journal of Medicine*, 381(10), 956–968. <https://doi.org/10.1056/NEJMr1806949>
- Sim, I. (2019b). *How you feel is important: Making PROs meaningful*. *Digital Healthcare Research: 2019 Year in Review*. Agency for Healthcare Research and Quality. <https://digital.ahrq.gov/2019-year-review/research-summary/emerging-innovative-newly-funded-research/how-you-feel-important-making-pros-meaningful>
- SleepyHead. (2018, April 26). *SleepyHead: Open Source CPAP Research and Review Software*. SourceForge. <https://sourceforge.net/projects/sleepyhead/>
- Smith–Lever Act of 1914. (2018). Chapter 79 of the 63rd Congress; Approved on May 8, 1914; 38 Stat. 372, 7 U.S.C. 341 et seq. Amended Through P.L. 115–334, Enacted December 20, 2018.

<https://www.govinfo.gov/content/pkg/COMPS-10296/pdf/COMPS-10296.pdf>

U.S. Copyright Office. (1998). *The Digital Millennium Copyright Act of 1998*.

<https://www.copyright.gov/legislation/dmca.pdf>

U.S. Copyright Office. (2015, October 28). *Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies*. Final rule.

<https://www.federalregister.gov/documents/2015/10/28/2015-27212/exemption-to-prohibition-on-circumvention-of-copyright-protection-systems-for-access-control>

U.S. Copyright Office. (2018, October 26). *Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies*. Final rule.

<https://www.federalregister.gov/documents/2018/10/26/2018-23241/exemption-to-prohibition-on-circumvention-of-copyright-protection-systems-for-access-control>

U.S. Copyright Office. (2021, October 28). *Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies*. Final rule.

<https://www.federalregister.gov/documents/2021/10/28/2021-23311/exemption-to-prohibition-on-circumvention-of-copyright-protection-systems-for-access-control>

Vogels, E. A. (2020, January 9). *About one-in-five Americans use a smart watch or fitness tracker*. Pew Research. <https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/>

Wallace, B. C., Schmid, C. H., Lau, J., & Trikalinos, T. A. (2009). Meta-Analyst: Software for meta-analysis of binary, continuous and diagnostic data. *BMC Medical Research Methodology*, 9, Article 80.

<https://doi.org/10.1186/1471-2288-9-80>

Wang, Y. (2020). *Meta-analysis of N-of-1 trials using Bayesian hierarchical models* [Doctoral dissertation, Brown University]. Biostatistics Theses and Dissertations. Brown Digital Repository. Brown University Library. <https://repository.library.brown.edu/studio/item/bdr:1129426/>

Wei, L. J., & Durham, S. (1978). The randomized play-the-winner rule in medical trials. *Journal of the American Statistical Association*, 73(364), 840–843. JSTOR. <https://www.jstor.org/stable/2286290>

Wheeler, M. (1976). *Lies, Damn lies, and statistics: The manipulation of public opinion in America*. Liveright.

Whitney, R. L., Ward, D. H., Marois, M. T., Schmid, C. H., Sim, I., & Kravitz, R. L. (2018). Patient perceptions of their own data in mHealth technology-enabled N-of-1 trials for chronic pain: Qualitative study. *JMIR mHealth and uHealth*, 6(10), Article e10291. <https://doi.org/10.2196/10291>

Wing, J. M. (2019). The data life cycle. *Harvard Data Science Review*, 1(1).

<https://doi.org/10.1162/99608f92.e26845b4>

Wolf, G. I., & De Groot, M. (2020). A conceptual framework for personal science. *Frontiers in Computer Science*, 2(21), 1–5. <https://doi.org/10.3389/fcomp.2020.00021>

World Health Organization. (2020, September 23). *Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation*. Joint statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC. <https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation>

Yao, Q., & Wei, L. J. (1996). Play the winner for phase II/III clinical trials. *Statistics in Medicine*, 15(21–22), 2413–2458. [https://doi.org/10.1002/\(sici\)1097-0258\(19961130\)15:22<2413::aid-sim460>3.0.co;2-n](https://doi.org/10.1002/(sici)1097-0258(19961130)15:22<2413::aid-sim460>3.0.co;2-n)

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J. A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalová, L., Pevsner-Fischer, M., Bikovsky, R., ... Segal, E. (2015). Personalized nutrition by prediction of glycemic responses. *Cell*, 163(5), 1079–1094. <https://doi.org/10.1016/j.cell.2015.11.001>

Zelen, M. (1969) Play the winner rule and the controlled clinical trial. *Journal of the American Statistical Association*, 64(325), 131–146. <https://doi.org/10.1080/01621459.1969.10500959>

Zucker, D. R., Schmid, C. H., McIntosh, M. W., D'Agostino, R. B., Selker, H. P., & Lau, J. (1997). Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *Journal of Clinical Epidemiology*, 50(4), 401–410. [https://doi.org/10.1016/s0895-4356\(96\)00429-5](https://doi.org/10.1016/s0895-4356(96)00429-5)

Zucker, D. R., Ruthazer, R., & Schmid, C. H. (2010). Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates: Methodologic considerations. *Journal of Clinical Epidemiology*, 63(12), 1312–1323. <https://doi.org/10.1016/j.jclinepi.2010.04.020>

Appendices

Appendix A. Should Data Scientists Be Portrayed as Gardeners/Farmers or Miners/Extractors?

The merits of ‘data gardening’ as a metaphor for personalized data science (Per-DS) are discussed in the main article, [Section 4.3](#), “Data Gardening versus Data Mining/Extraction.” We now elaborate on this discussion for data science in general, including both Per-DS and population-based data science (Pop-DS).

Data science is usually characterized as *the study of extracting value from data* (Columbia University Data Science Institute, 2020; Dominici & Parkes, 2021; Wing, 2019). Data science in general, and Pop-DS in particular, are often considered to be focused largely on data analytics, machine learning, and data mining, with less emphasis on study design and data collection. The use of the action word ‘*extracting*’ (as in ‘extracting value from data’) tends to reinforce the metaphor of ‘*mining*’ as the governing paradigm for data science.

Wing (2019) expanded this interpretation, explaining that ‘*extracting*’ represents the work done in all phases of the data life cycle, starting with the *generation* of data, followed by a sequence of collection, processing, storage, management, analysis, visualization, and interpretation. Meng (2019) concurred with Wing that “DS is not only about data analysis.” He also proposed an expansion of Wing’s framework: “the data generation process itself can entail several steps, such as goal setting, questionnaire or experimental design, and field testing.”

Dominici et al. (2021) further emphasized the importance of study design as a crucial component of data science, in particular, for estimating causal effects with observational data.

With the expanded scope for data science being “*not only about data analysis*” that was discussed above, we believe data scientists should avoid the confusion likely to result from the characterization ‘*extracting*,’ and instead characterize data science as being analogous to ‘*gardening*’ or ‘*farming*,’ to highlight the importance of the groundwork that lays the foundation for successful data science investigations, both in Per-DS and Pop-DS. Such groundwork (‘*cultivation*’) might entail identifying the goal for the investigation (‘goal setting’ as in Meng, 2019), identifying the data components needed, identifying and acquiring the sensors/devices/self-report instruments needed to generate/collect the data (‘questionnaire... design’), accessing data acquisition and analysis tools, and designing the study (when and how to deliver the intervention, when and how to collect the data; ‘experimental design, and field testing’). For Pop-DS, the groundwork might also include developing a partnership with the target population to facilitate community-based participatory research (CBPR) (Chung et al., 2006; National Institute on Minority Health and Health Disparities, 2018), which has the potential to foster long-term, mutually beneficial partnerships between data scientists and communities that house the studies. Such partnerships are widely known among real estate professionals as ‘*farming*.’ Maybe this ‘*farming*’ metaphor is something data scientists can learn from our colleagues in real estate, many of whom are experienced practitioners in data science.

As a side note, we opted to use the metaphor *data gardening* for Per-DS in the main article to highlight the home-based nature of Per-DS for personal knowledge cultivated, grown, and consumed by the individual Per-DS investigator. For Pop-DS, maybe *data farming* would be a more appropriate metaphor, without the connotation of being home-based/home-consumed, but still highlighting the importance of the groundwork to ‘*cultivate the field*’ in addition to the extraction/mining activities.

In summary, it might behoove the field to expand the characterization of data science from ‘*extracting value from data*’ to a more inclusive characterization such as ‘*gardening data for value*’ or ‘*farming data for value*,’ to highlight the need to include both ‘preparing the field’ to produce data of the greatest possible utility, and harvesting or mining or extracting value from the data thus produced, as essential components of data science.

Appendix B. PAP Therapy for Sleep Apnea: Fertile Field for Personalized Data Science (Per-DS)?

We now illustrate what personalized data science (Per-DS) can do by drawing upon the personal experience as one of us (ND) applied this methodology to inform his personal decisions using Positive Airway Pressure (PAP) therapy for sleep apnea. A brief synopsis was given in [Section 5.1](#). We now provide further details.

B.1. ND’s Personal Experience With PAP Therapy and Per-DS Investigations

ND was diagnosed with sleep apnea in 1999 and has used PAP therapy since. Almost every night he wears a mask connected through a hose to a PAP device that pumps air to create a small positive air pressure into his airway to keep it open, thereby reducing the occurrence of apnea/hypopnea events and snoring.

Over the years, ND has used approximately 10 PAP devices, beginning with a fixed pressure PAP device. Subsequently he switched to automatic PAP (APAP) devices with the capability to adjust the air pressure in real time to improve the quality of his sleep.

Air pressure setting is important for PAP therapy. Patients might not sleep comfortably if the pressure is too high for too long. Low pressure, on the other hand, might be inadequate to suppress apnea events. Skilled clinicians need to balance between the two extremes to find the optimal pressure setting for fixed-pressure PAP devices.

PAP devices record a rich variety of data in real time, including apnea indices, air pressure delivered, level of air leakage, and so on. The data can be transmitted via built-in cellular modem to the clinician for routine monitoring and intervention—the clinician can modify device setting remotely when needed.

APAP devices analyze the data collected in real time to predict upcoming apnea/hypopnea events and automatically adjust air pressure accordingly. When the device senses compromised airflow, it increases air pressure to open the airway to suppress the anticipated event, then decreases the pressure afterwards when normal airflow resumes. Such auto-titration helps deliver adequate pressure when needed while minimizing unnecessary long exposure to high pressure. Still, it is important for a skilled clinician to specify the maximum and minimum air pressures as limits for the range of air pressure auto-titration.

The data collected on the PAP device can be a valuable resource for Per-DS investigations. In recent years, device manufacturers have also provided online apps for patients to obtain summaries of the data transmitted from their devices to the manufacturer’s server, such as ResMed (2015).

Over the years, ND has conducted numerous Per-DS investigations on his PAP therapy. He has conducted systematic monitoring regularly on his therapeutic experience, using the key data elements recorded on his PAP device (apnea indices, air pressure, and air leakage), along with self-administered Epworth Sleepiness Scale (Johns 1991) for sleepiness during the day, and diary notes on the quality of his sleep each night. In recent years, he has also monitored his blood oxygen levels using an over-the-counter pulse oximeter connected to a smartphone, to provide continuous records of his blood oxygen level during his sleep. (He is aware of the caveat that over-the-counter oximeters are not of high quality for medical applications, therefore should be viewed as a crude marker at best.)

ND's systematic monitoring has mostly found his apnea indices and air leakage at a low level well within acceptable limits, indicating successful therapy. He finds this knowledge reassuring for his sustained adherence to the therapy. On some occasions, his monitoring detected some unusual events, such as unusually high air leakage, which usually indicated the need to adjust his headgear to tighten the seal, or the need to replace his mask. On some occasions, he observed somewhat elevated apnea indices at night and excessive sleepiness during the day, which he brought to the attention of his sleep medicine doctor, leading to discussions concerning adjustments to PAP device pressure settings or the need to replace the device.

Fortunately, ND has not encountered much need for trigger finding—the occasions when he had unusually high air leakage have always been resolved with headgear adjustment or mask replacement. There was a time when his 5-year old device was less effective in controlling apnea events. With the acquisition of a new device (covered by his insurance every 5 years), the problem was resolved.

ND also conducted a number of crossover trials with his PAP therapy configurations, mainly evaluating whether to switch to a new interface (oronasal mask, nasal mask, or nasal pillow). For example, several years ago he and his clinician considered a switch from an oronasal mask to a combination of a nasal mask and a chinstrap. ND used the oronasal mask to avoid massive air leakage due to his tendency to breathe with his mouth open during sleep. While the oronasal mask was a reasonable solution, it was bulky and uncomfortable. The less obtrusive nasal mask combined with a chinstrap might reduce mouth opening during sleep by holding the chin closed, leading to equal effectiveness and greater comfort. ND had bad experiences in the past with older versions of chinstraps that he found to be uncomfortable. However, this time the technician in the clinician's lab introduced him to a much wider chinstrap, imposing much less pressure on the chin.

ND's clinician recommended that he “try the two configurations back and forth for several weeks, to see which configuration feels better.” Therefore, ND conducted a personalized (N-of-1) trial, which informed his ultimate decision to switch to the nasal mask/chinstrap combination.

B.2. PAP Therapy Is Particularly Suitable for Per-DS Investigations

PAP therapy for sleep apnea is particularly suitable for Per-DS investigations for a number of reasons.

First, sleep apnea is a chronic condition with a substantial prevalence. Among adults aged 30–70 in the United States, the prevalence for mild to severe forms of this condition ($AHI \geq 5$) is estimated to be 26% (95% confidence interval: 24, 28), while the prevalence for moderate to severe condition is estimated to be 10% (95% confidence interval: 8, 11) (Peppard et al., 2013). PAP therapy is used widely to manage this condition. Many patients affected by this condition are likely to use PAP therapy for many years. Therefore, the lessons learned from a patient's Per-DS investigations have a good opportunity to deliver personalized benefits for years to come.

Second, the PAP industry continually improves the performance and convenience of PAP devices and accessories, such as reducing the size and weight and noise level of the devices, and making the interface (nasal mask, nasal pillow, and oronasal mask) more comfortable and effective. The life span for each product might only be a few years, and then new products are deployed with possibly better performance. Sleep apnea patients and their clinicians are therefore confronted with frequent opportunities, as well as challenges, to decide whether to continue using the current product configuration or to switch to a new one. Per-DS investigations, such as personalized trials, can be a useful tool to inform such decisions.

Third, there is substantial heterogeneity in patient responses to PAP therapy. A specific configuration of device, settings, and accessories might work well for one patient but not for another. Therefore, it is important for experienced sleep medicine clinicians to look beyond one-size-fit-all solutions, and customize PAP therapy to accommodate individual patients' unique needs, preferences, and capabilities (Watach et al., 2021). Per-DS investigations can be a useful tool to inform such decisions.

Fourth, there are already opportunities to personalize attributes for Per-DS investigations using data collected on PAP devices. For example, recent models of PAP devices produced by ResMed (2022) provide an option for the clinician to customize the patient menu on the device to specify what data elements are accessible to the patient on the device screen, with one option providing essential data only (mainly, the number of hours the device was used), and another option providing substantially more detailed data, including apnea indices, air pressure delivered, and level of air leakage. It is therefore possible for the clinician to personalize patient data access options according to the needs, preferences, and capabilities for each individual patient.

Beyond this simple option for the clinician to personalize patient data access on their PAP devices, it might be possible for manufacturers to provide additional options for the clinician to personalize patient interface with the PAP device to facilitate further Per-DS investigations, such as clinician-guided personalized trials for air pressure settings. Consider, for example, a scenario in which both the clinician and the patient are interested in exploring the possibility of increasing the maximum pressure on the patient's APAP device from the current setting of 12 cm H₂O to 14 cm H₂O, while leaving the minimum pressure at the current setting of 6 cm H₂O. Instead of making a one-time switch using the trial-and-error design, the clinician–patient team might find it more informative to conduct a personalized trial for several weeks, switching back and forth several times between the two maximum pressure settings to collect more data to evaluate the respective merits. In order to

conduct such a trial, the manufacturer would need to expand the current clinical menu on the device for the clinician to specify a range of the maximum pressures, such as 12 to 14 (instead of a single value such as 12),¹⁹ and authorize the patient to adjust the maximum pressure setting within this range in the patient menu on the device. With such an option, the clinician–patient team can easily set up a Per-DS investigation to make a more informed decision on the maximum pressure setting.

Fifth, ND’s insurance (an Advantage plan under Medicare) makes it relatively easy for him and his clinician to conduct personalized trials with new accessories such as new masks. He is allowed to return his accessories for a complimentary exchange with another product within 30 days in the event the accessory did not function satisfactorily. Patients whose insurance does not allow complimentary exchange might find it more difficult to conduct similar personalized trials.

A possible remedy might be for clinicians to be allowed to prescribe a combination order of accessories, say, 6 weeks of supply of mask A and 6 weeks of supply of mask B (instead of a 3-month supply of either mask) for a patient–clinician team interested in conducting a personalized trial, to decide whether to continue using mask A or to switch to mask B. Such a combination prescription, if allowed by the insurance carrier and the durable medical equipment (DME) supplier, will give patients and their clinicians more flexibility in reaching an evidence-based decision between the two masks.

Sixth, there is substantial demand from sleep apnea patients to better understand their PAP therapy and ways to improve their therapeutic experience and outcomes. One of us (DN), an experienced sleep medicine clinician, estimates that approximately 40% of sleep apnea patients in his clinical practice would be interested in some form of personal investigation using their own data if clinical and technical support were available. While DN’s practice might not be representative of all sleep apnea patients, it is plausible that a substantial proportion of sleep apnea patients overall have similar interests in Per-DS investigations.

Beyond ND’s personal experience and DN’s practice, there is a substantial interest among many sleep apnea patients to learn about their condition and take an active role in their PAP therapy. A number of data scientists have developed software, free to consumers, to access and analyze the data recorded on their PAP devices, such as SleepyHead (2018) (which is no longer supported by the developer) and Open Source CPAP Analysis Reporter (OSCAR) (Apnea Board 2019). There are extensive discussions among patients in bulletin boards such as Apnea Board (2011) on practical issues such as comparison between nasal mask versus nasal pillow, getting better control of air leakage, and so on. Many of those patients have explored informal trial-and-error Per-DS investigations. There is a rich potential for clinical scientists and data scientists to contribute further toward sleep apnea patients’ needs for support to their Per-DS investigations.

Finally, there is some evidence that sleep apnea patients who receive feedback from their own data will achieve better adherence with PAP therapy. For example, Malhotra et al. (2018) reported that PAP therapy patients who received active patient engagement (APE) via Resmed’s myAir app, which provided real-time feedback and

coaching to patients based on their own data collected from their PAP devices, achieved significantly a higher adherence rate than patients who received usual care without the APE intervention (87.3% vs. 70.4%; $P < .0001$). It is conceivable that wider applications of Per-DS investigations might achieve further improvements in patient engagement, adherence, and clinical outcomes.

Appendix C. Personalized Data Science (Per-DS) Through Millennia: Shennong, Barry Marshall, Daniel Carrión, and Beyond

In considering the prospect for *personalized data science* (Per-DS) to serve an important role in today's healthcare, we should not lose sight of history; we are not the first to have thought of careful self-investigation. This historical perspective is mentioned briefly in [Section 6](#). We now elaborate further on this perspective.

The ancient root for Per-DS can be traced back to the dawn of civilization, such as the legendary Emperor Shennong (the name means literally 'Divine Farmer'), who was alleged to have ruled China around 28th century BCE. According to the legends, Shennong tasted hundreds of herbs to study their medicinal effects on his body. The compendium attributed to Shennong,²⁰ "*Shén Nóng Běncǎo Jīng: The Divine Farmer's Classic of Materia Medica*" (Wilms, 2017/ca. 201–300 CE), documents the medicinal effects of hundreds of herbals.

According to legends, Shennong's body was transparent; therefore, he was able to visualize how his body reacted to the herbal medicine being tested. This is an inspiring recognition of the important role of data collection for Per-DS, if not by Shennong in reality, then by storytellers who passed on Shennong's legends.

Modern medical technologies such as continuous glucose monitors, pulse oximeters, smart watches with electrocardiogram capability, positive airway pressure (PAP) devices, and sensitive thermometers for monitoring daily temperature are analogous to Shennong's legendary transparent body, allowing Per-DS investigators to 'see' their own glucose levels, oxygen levels (both for sleep apnea and for monitoring COVID progression), heart rhythms, sleep patterns, and ovulation status in digital form, making the invisible, visible. Such technologies enable individuals to conduct Per-DS studies to inform their personal health decisions.

It is a remarkable coincidence that Emperor Shennong not only pioneered Per-DS investigations in herbal medicine, but also led the development of agriculture in ancient China. While we honor the synergy between farming and medicine in Shennong's legends, we also note a similar connection between data gardening/data farming and data science, as discussed in [Section 4.3](#) and Appendix A, "Should Data Scientists Be Portrayed as Gardeners/ Farmers or Miners/Extractors?"

In modern times, Per-DS was practiced extensively by clinical investigators conducting self-experiments on themselves, with major accomplishments including seven Nobel Prizes (Altman, 1987; Hanley, et al., 2019; Weisse, 2012). For example, Barry J. Marshall's research on *Helicobacter pylori* and peptic ulcer disease won him and his collaborator J. Robin Warren the Nobel Prize in Physiology or Medicine in 2005, "for their discovery of the bacterium *Helicobacter pylori* and its role in gastritis and peptic ulcer disease." (The Nobel

Prize, 2022) This award was based largely on Marshall’s self-experiment in which he ingested a broth containing cultured *H. pylori*, along with endoscopies taken both before and after showing the development of gastritis and biopsy showing colonization of *H. pylori* in his stomach, providing strong evidence that ingested *H. pylori* was the cause for the gastritis.

While many Per-DS investigations practiced by clinical investigators have succeeded in contributing toward medicine throughout the millennia, not all Per-DS investigations have a happy ending. Shennong’s legend came to a tragic conclusion when he ingested the yellow flower of a weed that caused his intestines to rupture, killing him before he could take his antidote tea.

Shennong was not alone in sacrificing his life for Per-DS investigations. Daniel Carrión, a Peruvian medical student, conducted a fatal self-experiment in 1885 when he was inoculated with blood taken from a patient infected with the disease now known as Carrión's disease. The experiment led to his death but also proved the causal relationship between the acute and chronic forms of the disease. Peru declared him a “National Hero, Martyr and Master of Medicine”; the day he died from the disease, October 5, was declared the “Day of Peruvian Medicine.” (Andean News Agency, 2021; Carrasco, 2020)

Without diminishing the honor owed to history’s Per-DS heroes, their accomplishments were often not without personal sacrifice. Today’s lay Per-DS investigators can learn from history—while embracing the potential benefits from their self-investigations, they need to seek competent professional guidance from both clinicians and data scientists before taking on a self-investigation with possible risks for self-harm. Further discussions on the delivery of such professional guidance virtually are given in [Section 5.2](#).

Appendices References

Altman, L. K. (1987). *Who goes first? The story of self-experimentation in medicine*. Random House.

Andean News Agency. (2021, October 5). Peru's President attends ceremony marking Day of Peruvian Medicine. <https://andina.pe/ingles/noticia-perus-president-attends-ceremony-marking-day-of-peruvian-medicine-864312.aspx>

Apnea Board. (April 23, 2019). *OSCAR - Open Source CPAP Analysis Reporter*. 04-23-2019. <http://www.apneaboard.com/forums/Thread-OSCAR-Open-Source-CPAP-Analysis-Reporter--25239>.

Apnea Board. (2011, March 5). *Apnea Board: Sleep apnea patients helping one another.*, from <https://www.apneaboard.com/>

Carrasco, J. L. A. S. (2020, October 5). Día de la Medicina Peruana. Escuela Nacional de Salud Pública. [Day of Peruvian Medicine. National School of Public Health.] http://www.minsa.gob.pe/ensap/comunicados/2020/comunicado_78_2020.pdf

Chung, B., Corbett, C. E., Boulet, B., Cummings, J. R., Paxton, K., McDaniel, S., Mercier, S. O., Franklin, C., Mercier, E., Jones, L., Collins, B. E., Koegel, P., Duan, N., Wells, K. B., Glik, D., & Talking Wellness Group of Witness for Wellness. (2006). Talking Wellness: A description of a community-academic partnered project to engage an African-American community around depression through the use of poetry, film, and photography. *Ethnicity & Disease*, 16(1 Suppl 1), S67–S78. PMID: 16681130. <https://pubmed.ncbi.nlm.nih.gov/16681130/>

Columbia University Data Science Institute. (2020, September 22). *About DSI: Data science is the study of extracting value from data*. Retrieved January 3, 2022, from <https://datascience.columbia.edu/about-us/>

Dominici, F., Bargagli-Stoffi, F. J., & Mealli, F. (2021). From controlled to undisciplined data: Estimating causal effects in the era of data science using a potential outcome framework. *Harvard Data Science Review*, 3(3). <https://doi.org/10.1162/99608f92.8102afed>

Dominici, F., & Parkes, D. C. (2021). Introducing the interim co-editors-in-chief. *Harvard Data Science Review*, 3(3). <https://doi.org/10.1162/99608f92.24ee17a7>

Hanley, B. P., Bains, W., & Church, G. (2019). Review of scientific self-experimentation: Ethics history, regulation, scenarios, and views among ethics committees and prominent scientists. *Rejuvenation Research*, 22(1), 31–42. <https://doi.org/10.1089/rej.2018.2059>

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep*, 14(6), 540–545. <https://doi.org/10.1093/sleep/14.6.540>

Malhotra, A., Crocker, M. E., Willes, L., Kelly, C., Lynch, S., & Benjafield, A. V. (2018). Patient engagement using new technology to improve adherence to positive airway pressure therapy: A retrospective analysis. *Chest*, 153(4), 843–850. <https://doi.org/10.1016/j.chest.2017.11.005>

Meng, X.-L. (2019). Data science: An artificial ecosystem. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.ba20f892>

National Institute on Minority Health and Health Disparities. (2018). *Community-Based Participatory Research Program (CBPR)*. Updated October 2, 2018. <https://www.nimhd.nih.gov/programs/extramural/community-based-participatory.html>

Peppard, P. E., Young, T., Barnet, J. H., Palta, M., Hagen, E. W., & Hla, K. M. (2013). Increased prevalence of sleep-disordered breathing in adults. *American Journal of Epidemiology*, 177(9), 1006–1014. <https://doi.org/10.1093/aje/kws342>

ResMed. (2015, May 4). *MyAir*. <https://myair2.resmed.com/>

ResMed. (2022, August 2). *ResMed*. <https://www.resmed.com/en-us/>

SleepyHead. (2018, April 26). *SleepyHead: Open Source CPAP Research and Review Software*.
<https://sourceforge.net/projects/sleepyhead/>

The Nobel Prize. (2022, August 26). The Nobel Prize in Physiology or Medicine 2005. Nobel Prize Outreach AB 2022. <https://www.nobelprize.org/prizes/medicine/2005/summary/>

Watach, A. J., Hwang, D., & Sawyer, A. M. (2021). Personalized and patient-centered strategies to improve positive airway pressure adherence in patients with obstructive sleep apnea. *Patient Preference and Adherence*, 15, 1557–1570. <https://doi.org/10.2147/PPA.S264927>

Weisse A. B. (2012). Self-experimentation and its role in medical research. *Texas Heart Institute Journal*, 39(1), 51–54. <https://pubmed.ncbi.nlm.nih.gov/22412227/>

Wilms, S. (Trans.) (2017). *Shén Nóng Běncǎo Jīng: The Divine Farmer's Classic of Materia Medica* (3rd ed.). Happy Goat Productions. <https://www.happygoatproductions.com/onlinestore/the-divine-farmers-classic-of-materia-medica-3rd-edition>. (Original work published ca. 201–300 CE).

Wing, J. M. (2019). The data life cycle. *Harvard Data Science Review*, 1(1).
<https://doi.org/10.1162/99608f92.e26845b4>

©2022 Naihua Duan, Daniel Norman, Christopher Schmid, Ida Sim, and Richard L. Kravitz. This article is licensed under a Creative Commons Attribution (CC BY 4.0) [International license](https://creativecommons.org/licenses/by/4.0/), except where otherwise indicated with respect to particular material included in the article.

Footnotes

1. We use the term ‘population’ loosely here, as the collection of data donors might not be representative of a target population. For some Pop-DS applications, the collection of individuals might be self-representing. For example, when a company studies its customers’ purchase behavior to inform future customer service decisions, the target population coincides with the company’s customer pool of data donors. However, some Pop-DS applications might attempt to generalize to potential clients outside the existing client pool, therefore the collection of data donors might not be representative of the entire target population. [↵](#)
2. While this article addresses both health and wellness, we will sometimes abbreviate the reference to health, implying that wellness is included in a broad conceptualization of health. [↵](#)
3. In order to balance the traditional dominance of male gender in research literature, we use female gender pronouns as representative of both genders. [↵](#)
4. The Per-DS investigator is the solo ‘participant’ in her own study, not a ‘participant’ in the usual sense of being one of many participants in a study directed by someone else such as a principal investigator. [↵](#)

5. We use the terms ‘patient’ and ‘individual’ interchangeably referring to Per-DS investigators who study their own health-related data to inform their own health-related decisions. Those might be patients with disease conditions seeking to improve their treatment decisions, or healthy individuals seeking to improve their health and wellbeing. [↵](#)
6. None of the displays presented to the interviewees in Whitney et al. (2018) was in a tabular form, therefore this study does not inform the size of the subpopulation who prefer tabular presentations to graphical presentations. While the majority of end-users probably prefer graphs to tables, the preference of the minority who prefer tables to graphs still deserve to be accommodated in a personalized world. [↵](#)
7. Here ‘treatment’ is meant to include both medical and behavioral interventions, and might include ‘no active treatment’ as one of the options to be evaluated. [↵](#)
8. With an assignment sequence like ABBA, the deterioration between the first week (treatment A) and the second week (treatment B) is offset by the deterioration between the third week (treatment B) and the fourth week (treatment A), so the overall comparison between the A weeks and the B weeks is not confounded with a linear deterioration. Assignment sequences like ABAB are confounded with the linear deterioration: the deterioration between the first A week and the first B week, and the deterioration between the second A week and the second B week, are accumulated in the comparison between the A weeks and the B weeks. [↵](#)
9. Being ‘stable’ includes conditions that fluctuate randomly within a limited range without a time trend. [↵](#)
10. We assume implicitly that the phenomenon being studied is stable over time, therefore, personal knowledge gained from ‘the patient that is me today’ can be applied to ‘the patient that will be me tomorrow.’ [↵](#)
11. We appreciate Deborah Estrin’s insightful suggestion to use the characterization ‘gardening’ instead of ‘farming’ as being the most quintessential to Per-DS. Per-DS is like gardening in being home-based and amateurish, and can be practiced by those willing in the general population rather than restricted to experts, with limitations such as avoiding unsafe self-experimentations, analogous to avoiding toxic plants in home gardening. The lay nature of Per-DS under the gardening metaphor does not necessarily predicate poor quality, to wit, those who profess homegrown beefsteak tomatoes as being unmatched in taste compared to commercially produced tomatoes. The metaphor ‘farming’ captures some of the essence of Per-DS, such as the need to cultivate the field. However, farm products are usually marketed to others for their consumption; therefore, ‘farming’ does not quite capture the focus for Per-DS to produce personal knowledge for self-consumption. [↵](#)
12. It might be possible to combine personal data produced in Per-DS investigations and synthesize/meta-analyze combined data to produce generalizable knowledge that can be applied to benefit others. Under this

scenario, the investigation is no longer a ‘pure’ Per-DS investigation, but is rather a hybrid between Per-DS and Pop-DS, and human subjects protection regulations might be applicable. [↵](#)

13. We refer to Per-DS studies as *investigations* instead of *research*, to clarify the distinction between personal knowledge and generalizable knowledge. [↵](#)

14. The driver who holds the steering wheel still needs to submit to oversight by regulatory agents such as traffic police. Nevertheless, the driver is the one who determines the destiny for the vehicle. [↵](#)

15. The CDC estimated that 19.6 million (8.0% of) U.S. adults had high-impact chronic pain in 2016 (Dahlhamer et al., 2018). We believe it is reasonable to assume that at least half of high-impact chronic pain patients suffer chronic musculoskeletal pain, thus there are 9.8 million high-impact chronic musculoskeletal pain patients. [↵](#)

16. $9.8 \text{ million} \times 33.0\% \times 59.7\% = 1.9 \text{ million}$. [↵](#)

17. The prevalence for moderate-to-severe sleep apnea is estimated to be 10% (95% confidence interval: 8,11) among adults aged 30–70 in the United States (Peppard et al. 2013). Based on the U.S. Census (Howden & Meyer 2011), there are 154,957,413 adults aged 30–69 in the United States in 2010, yielding an estimate of 15,495,741 moderate-to-severe sleep apnea patients. Taking DN’s estimate of 40% yields an estimate of 6,198,296 prospective Per-DS investigators. [↵](#)

18. After a candidate trigger or deterrent is identified, the investigation might progress into a confirmation phase to confirm experimentally the effect of the candidate, for example, using a pre-post trial or a personalized trial. The risk level for those investigations, and the need for clinician involvement, would be higher than during the observational phase of the trigger finding investigation. [↵](#)

19. Even with such an option enabled by the manufacturer in the clinical menu, the clinician can easily disable patients from conducting unauthorized personalized trials on the maximum pressure setting by specifying the same value (such as 12 and 12) for the range of maximum pressure, instead of a genuine range (such as 12 and 14) to authorize the patient to conduct a personalized trial. [↵](#)

20. This compendium is believed to be written around the second century CE by an anonymous author. [↵](#)