DTC-Genetics, the quantified self and patient-driven research

03.01.12, Bastian Greshake

some words about me

- Bachelor in Life Sciences (2010)

- Scholarship/Research Assistant at Biodiversity & Climate Research Center (since 2010)

- Master studies at the Goethe University in Frankfurt/Main (since 2011)

- Not exactly a biologist with much professional background in human genetics, but...

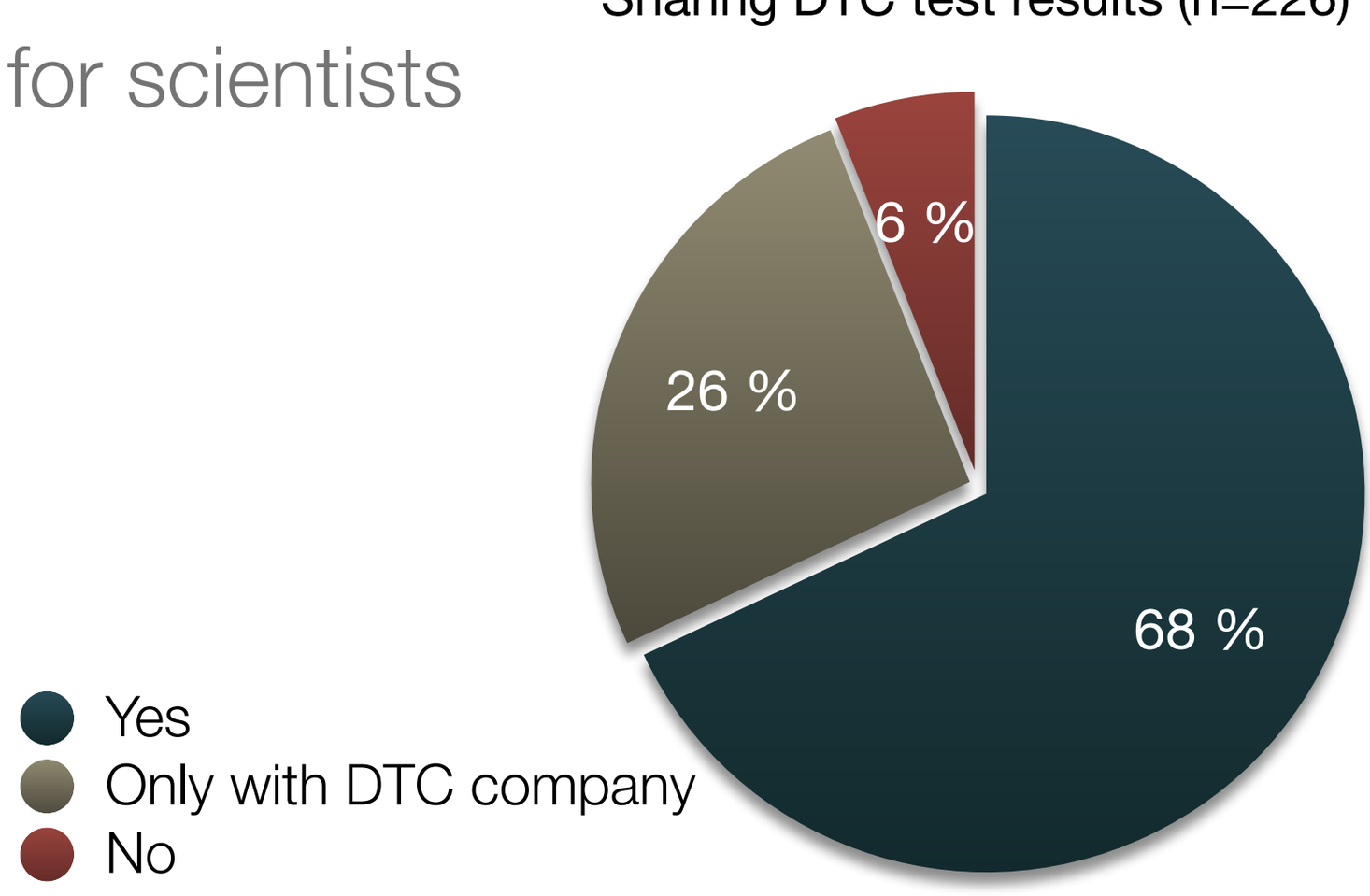
some words about me

- customer of DTC genetic testing
- some background in data mining
- some experience with web application design
- interest in social media & crowd-sourcing

```ruby
def self.perform(snp_id)
   @snp = Snp.find(snp_id)
   if @snp.mendeley_updated < 31.days.ago
     key_handle = File.open(::Rails.root.to_s+"/key_mendeley.txt")
     api_key = key_handle.readline.rstrip

     url = "http://api.mendeley.com/oapi/documents/search/"+@snp.name+"/?c

     begin
        resp = Net::HTTP.get_response(URI.parse(url))
     rescue
        retry
     end

     data = resp.body
     result = JSON.parse(data)

     if result["total_results"] != 0
        print "mendeley: Got papers\n"
        result["documents"].each do |document|
           mendeley_url = document["mendeley_url"]
           uuid = document["uuid"].to_s
           first_author = document["authors"][0]["forename"]+" "+document[
           title = document["title"]
           pub_year = document["year"]
           doi = document["doi"]

           if MendeleyPaper.find_all_by_uuid(uuid) == []
              print "-> paper is new\n"
              @mendeley_paper = MendeleyPaper.new(:mendeley_url => mendeley
              if doi != []
                 @mendeley_paper.doi = doi
              end

              @mendeley_paper.save
              @snp.ranking = @snp.mendeley_paper.count + 2*@snp.plos_paper.

              print "-> Written paper\n"
           else
              print "-> paper is old\n"
              @mendeley_paper = MendeleyPaper.find_by_uuid(uuid)
           end
           Resque.enqueue(MendeleyDetails,@mendeley_paper)
        end
     else
        print "mendeley: No papers found\n"
     end
     @snp.mendeley_updated = Time.zone.now
     @snp.save
     sleep(1)
```

mining DTC genetic tests

- lots of potential for open data (100k+ customers)
- people are willing to share (and are doing so)
- cheap data source for scientists
- but?



mining DTC genetic tests

- people need motivation to share this kind of information
- plus a low entry barrier to participate
- and the data should be re-usable and needs to be searchable (findable!)

Google Docs

public 23nM datasets Go to spreadsheet view – Hide colors Print...

IND PGP BuzzData openSNP

Row ▲		Name / username (or not)	url (this spreadsheet URL is: http://bit.ly/cNDHf3)
	+ Add	(All)	(All)
1	Edit	V. Alex Brennen / vab	http://vab.mit.edu/
2	Edit	Hong ChangBum / 23nM: hongiiv	http://www.snpedia.com/index.php/User:Hongiiv
3	Edit	Mark Davis / m86	http://www.markcoenterprises.net/files/23andMe
4	Edit	Gene Girard	http://www.snpedia.com/index.php/User:Gene_Girard
5	Edit	Mikolaj Habryn / dichro / 23nM: Mikolaj	https://sites.google.com/a/rcpt.to/dichro/Home/genotyping
6	Edit	Joakim Jardenberg / 23nM: jardenberg	http://jardenberg.se/b/min-dna-profil-projekt-transparens/
7	Edit	John Kemeny / 23nM: john26	https://www.23andme.com/you/community/thread/4861/
8	Edit	Hyungyong Kim / yong27	http://biohackers.net/wiki/yong27/Genome
9	Edit	James I. McMillian / 23nM: BlueEyesRedHair	http://www.box.net/shared/edlj33z0rz
10	Edit	Andrew Morrow / 23nM: AndrewMorrow	http://amorrow2.pbworks.com/
11	Edit	Paul Shenton / 23nM: Pupsenok	https://www.23andme.com/user/?community_profile=b62913e07f500
12	Edit	Sam Snyder	http://www.snpedia.com/index.php/User:Samsnyder
13	Edit	Mike Spear / 23nM: Mikes Gene	http://www.genomealberta.ca/personalgenotyping/default.aspx
14	Edit	Elise Wood / glytch	http://glytch.org/blog/posts/2009_03_18_my_genome_from_23andme
15	Edit	Debra Patek / 23nM: familypast	http://www.familypast.com/genome_familypast_Full_20110301083300
16	Edit	Michael Lally / 23nM: Maolalai	http://www.box.net/shared/tzym2lsik6
17	Edit	Grant Brunner	http://www.archive.org/details/GrantBrunners23andmeResults012511
18	Edit	Manu Sporny	http://manu.sporny.org/2011/public-domain-genome/
19	Edit	orta	http://github.com/orta/dna
20	Edit	Manuel Corpas and family (multiple)	http://manuelcorpas.com/five-family-relatives-genome-download/
21	Edit	Kenneth Reitz	https://github.com/kennethreitz/genome
22	Edit	Razib Khan	http://razib.com/razib.zip
23	Edit	Ramūnas Janavičius	http://cancergenetics.wordpress.com/2011/04/13/my-genome-scan-m
24	Edit	Beau Gunderson	https://github.com/beaugunderson/genome
25	Edit	Bastian Greshake	https://github.com/gedankenstuecke/genotyping_23andme
26	Edit	Genomes Unzipped (12 individuals)	http://www.genomesunzipped.org/data
27	Edit	Nils Homer	http://nilshomer.com/index.php?title=My_Genetic_Data
28	Edit	anonymous, Ashkenazi	https://rapidshare.com/files/961006780/ASHK.tar.gz
29	Edit	Sebastian Bassi	https://github.com/sbassi/MiGenomaSbassi

Google public genomes

Alles Bilder Videos News

NOVA | **Public Genomes** - [Diese Seite übersetzen]
www.pbs.org/wgbh/nova/body/**public-genomes**.html
18 Aug 2009 – Many people are paying to have their **genomes** on the Internet. Are they crazy?

SNPedia

Page Discussion

Genomes

This spreadsheet of genomes with raw

Navigation

SNPedia
Promethease
FAQ
Blog
Current events
Recent changes
Random page

These are the 104 public genomes. They genomes. Nearly all of them have regular to date each is.

	Platform	
Aaron Vollrath	23andMe v2	
Almelina	23andMe v2	1

gedankenstuecke / **genotyping_23andme**

Code Network

My raw genotyping data from 23andme — Read more

ZIP HTTP Git Read-Only https://github.com/gedanke

Files Commits Branches 1 Tags Downloac

Latest commit to the **master** branch

edit on readme

finding DTC results up to now

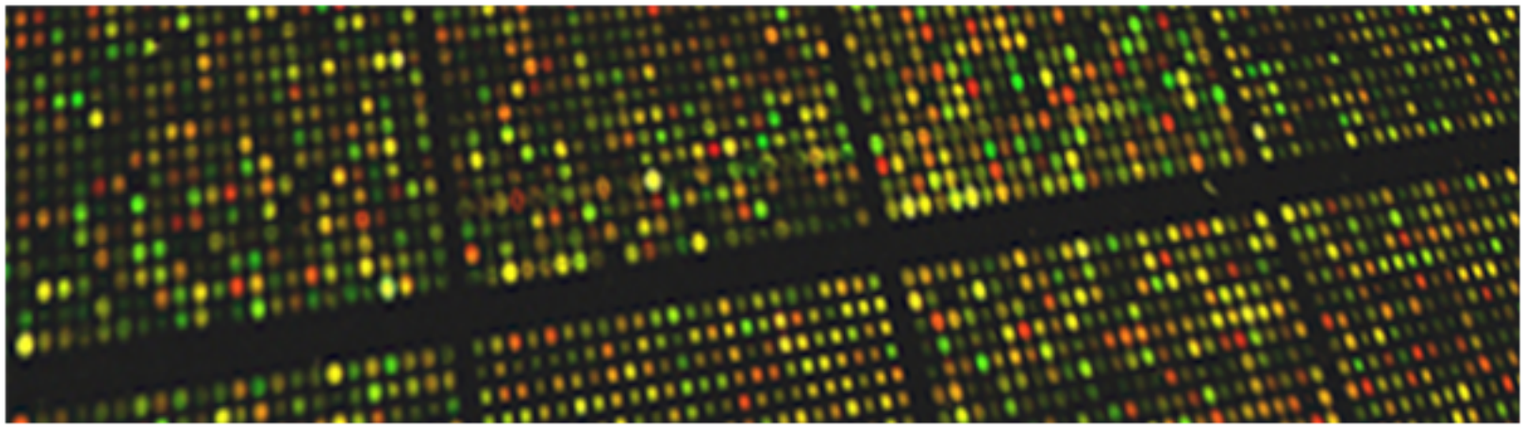
mining DTC genetic tests

- results are hidden somewhere on the web
- often no phenotypic annotation
- not easily re-usable

openSNP

- wants to be a central repository for sharing DTC results
- enables users to share phenotypes as well
- lowers barrier to participate
- motivation via benefits for users
- ultimate aim: enabling GWAS

Welcome to *openSNP*



openSNP **allows customers of direct-to-customer genetic tests to publish their test results, find others with similar genetic variations, learn more about their results, find the latest primary literature on their variations and help scientists to find new associations.**

Sign Up!

For Genotyping Users | For Scientists | FAQ

Upload Your Genotyping File

Upload the genotyping raw-data you got from 23andMe or deCODEme to the database of *openSNP* to make it available for everybody.

Share Your Phenotypes & Traits

Share as many phenotypes, characteristics and traits with other *openSNP* users and find others with similar characteristics.
And maybe help scientists to discover new genetic associations.

Share your stories on variations & phenotypes

openSNP lets you share your stories on your genetic variations & phenotypes with others. Discover the stories of other users.
Find others to exchange experiences about your variations.

Find literature on genetic variation

openSNP gets the latest open access journal articles on genetic variations via the Public Library of Science. Additionally popular articles are indexed via the social reference manager Mendeley. Summaries are provided by SNPedia.

the front

Hello Bastian!

Description of yourself

Life Scientist, currently studying ecology and evolution in Frankfurt/Main, Germany and one of the founders of openSNP. Feel free to message me if you encounter bugs.

Variations you did not enter yet (13) | Your variations | Your messages | Replies to your comments

Name of Phenotype	Created at
Sex	27.12.2011 15:14
Faktor 5 Leiden (F5)	27.12.2011 13:40
Number of great-great grandparents born in Sweden	06.12.2011 18:43
Brown hair, brown eyes, pale skin. B+ blood type. Central Asian	03.12.2011 07:40
Kell Blood Group (K/k antigens)	03.12.2011 01:33
Hypertriglyceridemia	03.12.2011 01:11
Diego Blood Group	03.12.2011 00:58
ENTP	02.12.2011 12:33
SAT Writing	02.12.2011 00:31

Phenotype: Number of toes

Back to your homepage
Back to your phenotypes

KNOWN PHENOTYPES:

- **10 (5 + 5), Other than 10 (5 + 5)** Download genotyping-files of all corresponding users
- **10 (5 + 5)** Download genotyping-files of all corresponding users

Description | Users sharing this phenotype (4) | Comments (0)

User	Variation
Beau Gunderson	10 (5 + 5)
Mark Davis	10 (5 + 5)
Martin	10 (5 + 5)
Bastian Greshake	10 (5 + 5)

other resources

- Personal Genome Project
    - data is open
    - participation not

Personal Genome Project

Home | Project Overview | Participation Overview | PGP Community | DONATE

Volunteers from the general public working together with researchers to advance personal genomics.

We believe individuals from the general public have a vital role to play in making personal genomes useful. We are recruiting volunteers who are willing to share their genome sequence and many types of personal information with the research community and the general public, so that together we will be better able to advance our understanding of genetic and environmental contributions to human traits. Learn more about how to **participate** in the Personal Genome Project.

Project Overview. The PGP hopes to make personal genome sequencing more affordable, accessible, and useful for humankind. Learn more about our **mission**.

Want to participate? We aim to enroll 100,000 informed participants from the general public. Learn more about **participation** in the PGP and how you can get involved.

Meet our volunteers. Participants may volunteer to publicly share their DNA sequence and other personal information for research and education. Meet the "**PGP-1K**".

Participant Login

Login Now

Project News

Subscribe to our newsletter.

Oct 5, 2011:PGP-HMS prepares for national blood collection campaign, adds hundreds of walk-in clinics to network.**See list**.

Sep 10, 2011:KPGP publishes 32 genomes of Korean participants. **More**.

Personal Genome Project

other resources

- Personal Genome Project
    - data is open
    - participation not
- genomera
    - participation is open
    - focus on small scale studies/experiments

genomera Heal the world.

Groups Studies Instruments Profile gedankenstuecke

Processing Reality: Impact of Dopamine Modulation on Memory Filtering

IN DESIGN

You are a Data Participant

Invite Friends | Change Role

Objective: To determine if genetic variants related to dopamine processing in the brain impact the processing of memories according to their relation with ongoing reality

Feedback

What would you like to contribute?

gedankenstuecke joined as a data participant.

less than a minute ago

Comment...

lwbaum joined as a data participant.

about 11 hours ago

Organizers

cwhogg

melanie

Participant Tasks

1) Complete background demographic survey(s) (10 min)
2) Complete memory filtering tasks (2x) (30 min)

Participant Requirements

1) Absence of psychological or neurologic disorders (ie. bipolar disorder, schizophrenia, epilepsy, parkinson's, prior stroke, traumatic brain injury, dementia)

genomera

genomera

- started out of DIYGenomics
- for users: Find interesting studies to participate in
- enables citizen science
- not limited to association studies
- no results so far

more patient driven research

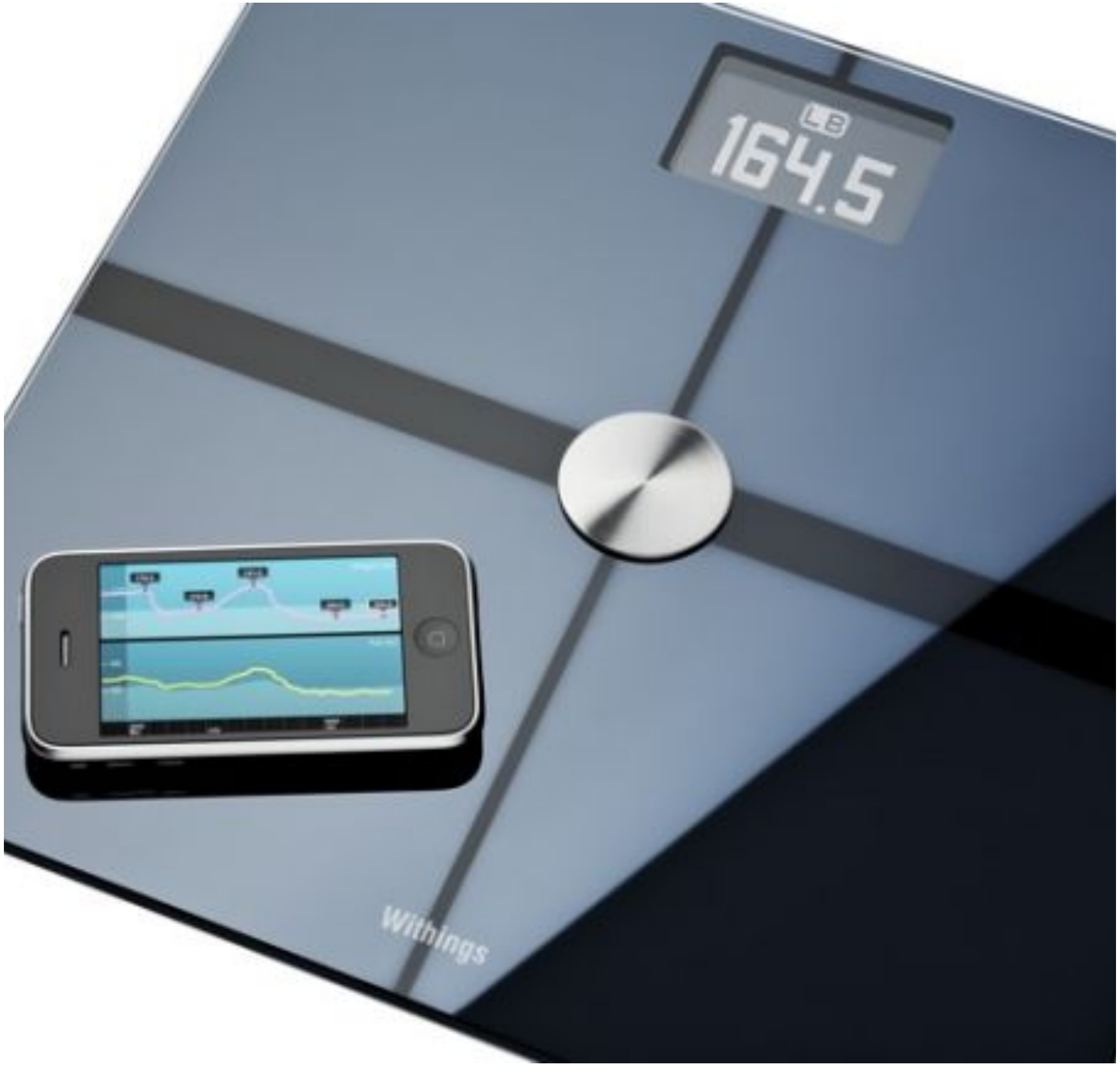
- PatientsLikeMe
    - around since ~2006
    - published a dozen studies since then
    - famous example: ALS research on lithium intake

problems of patient driven research

- sample sizes
- bias in participants
- no blind experiments
- motivation of participants
- accuracy of data

potential of patient driven research

- possible sample sizes
- low costs
- "warm fuzzy feeling inside" for patients



Quantified Self Movement

quantified self

- new data sources
    - eating habits
    - work-outs
    - sleep habits
    - ...
- automated through technology
    - smart phones

Quantified Self and Science

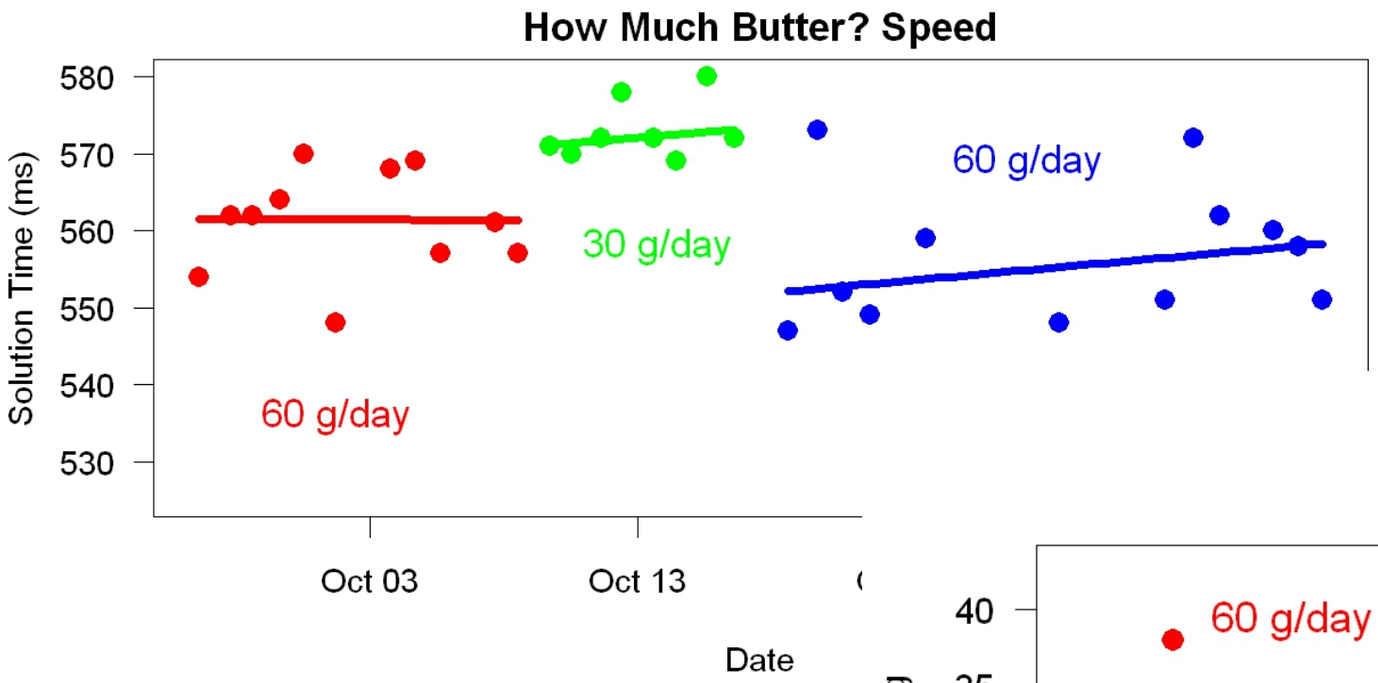
QS projects

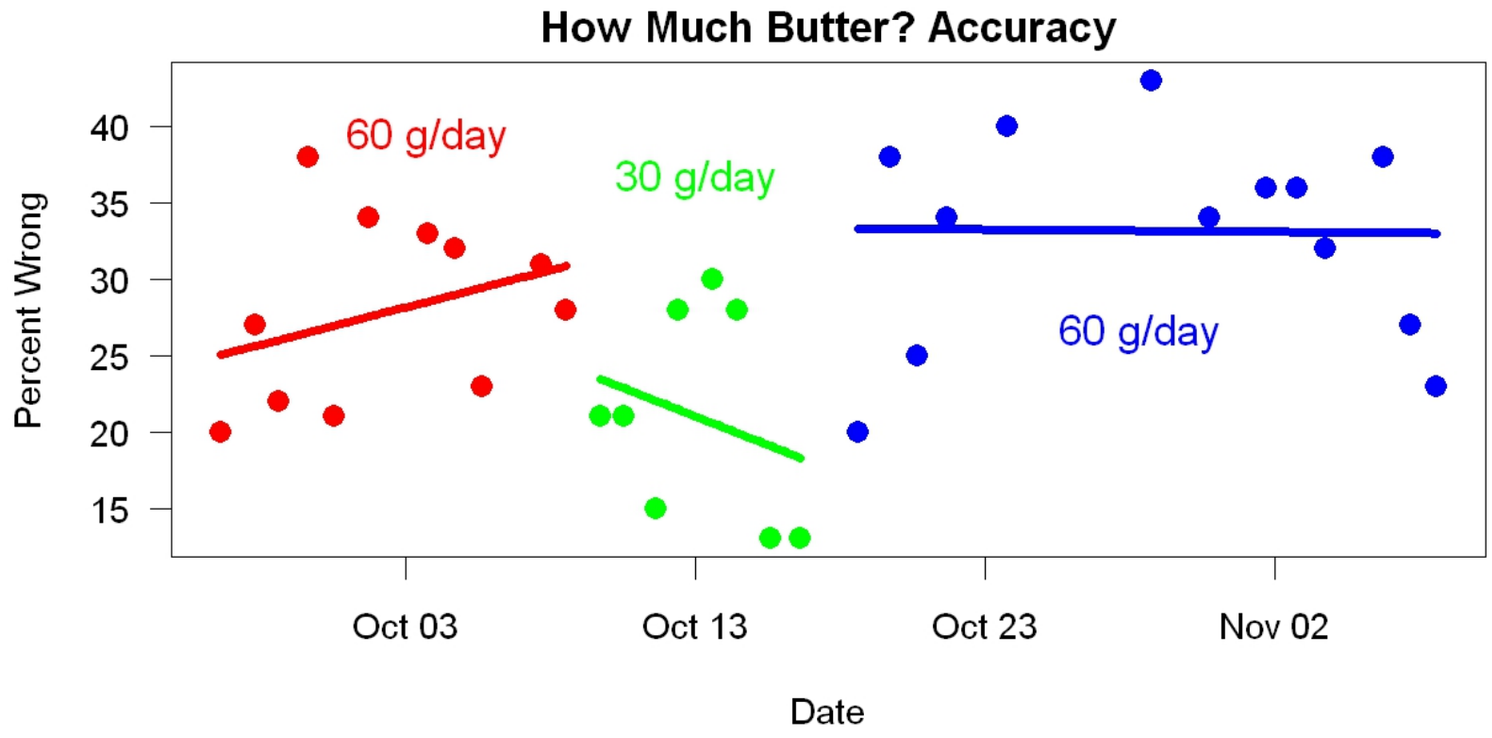
- tracking health in response to work-outs (minimizing impacts of disease/genetic predisposition)

- track response to different drugs

- tracking well-being in response to eating habits (butter vs arithmetics)

butter vs arithmetics





source: Seth Roberts - quantifiedself.com

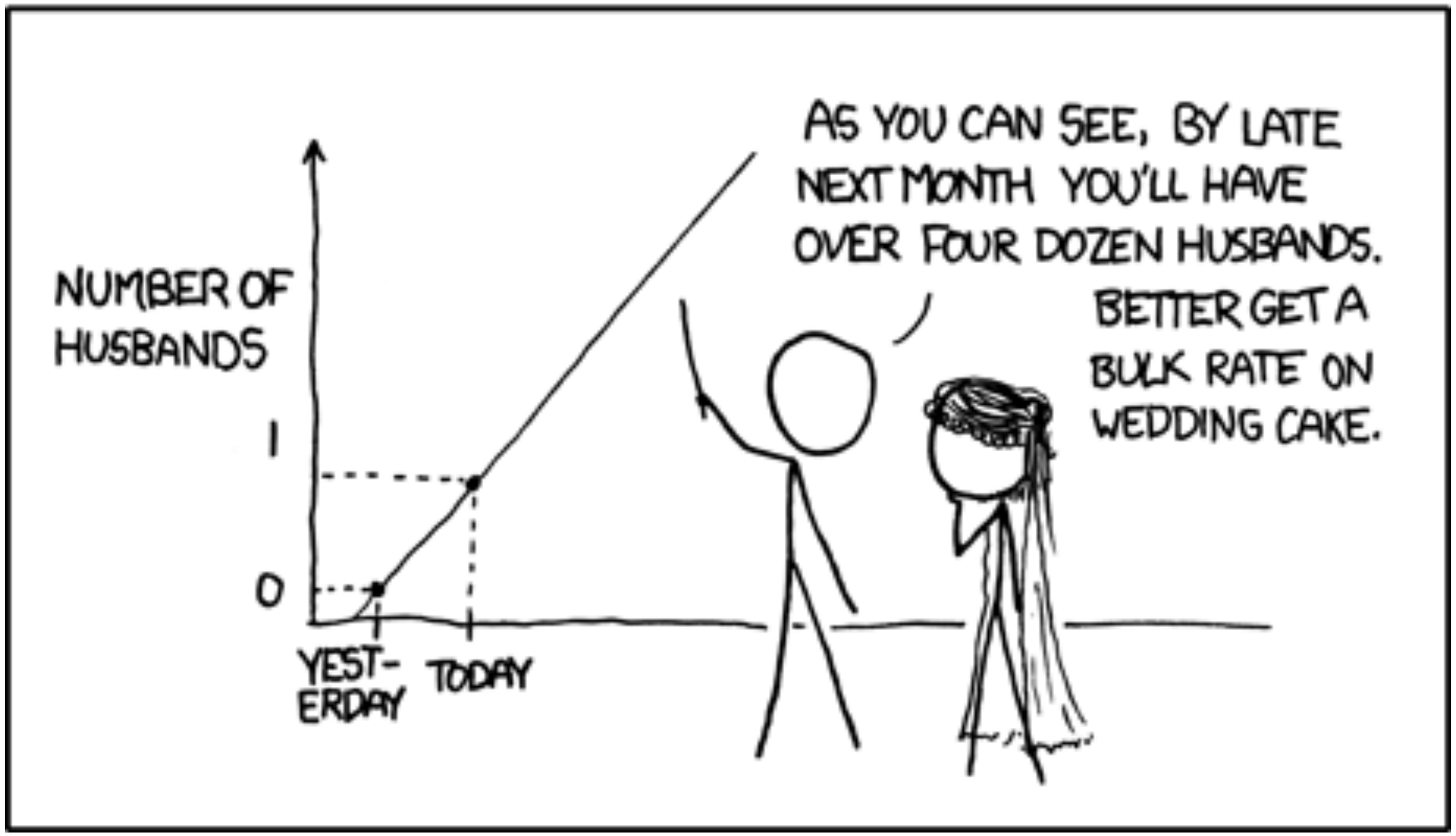
Conclusions

- DTC results and patient driven research can lead to new scientific knowledge

- can be performed in addition to traditional research

- technology enables new kinds of research

- can we include the Quantified Self Movement in our studies?



thanks for your attention

time to discuss this

source: xkcd.com