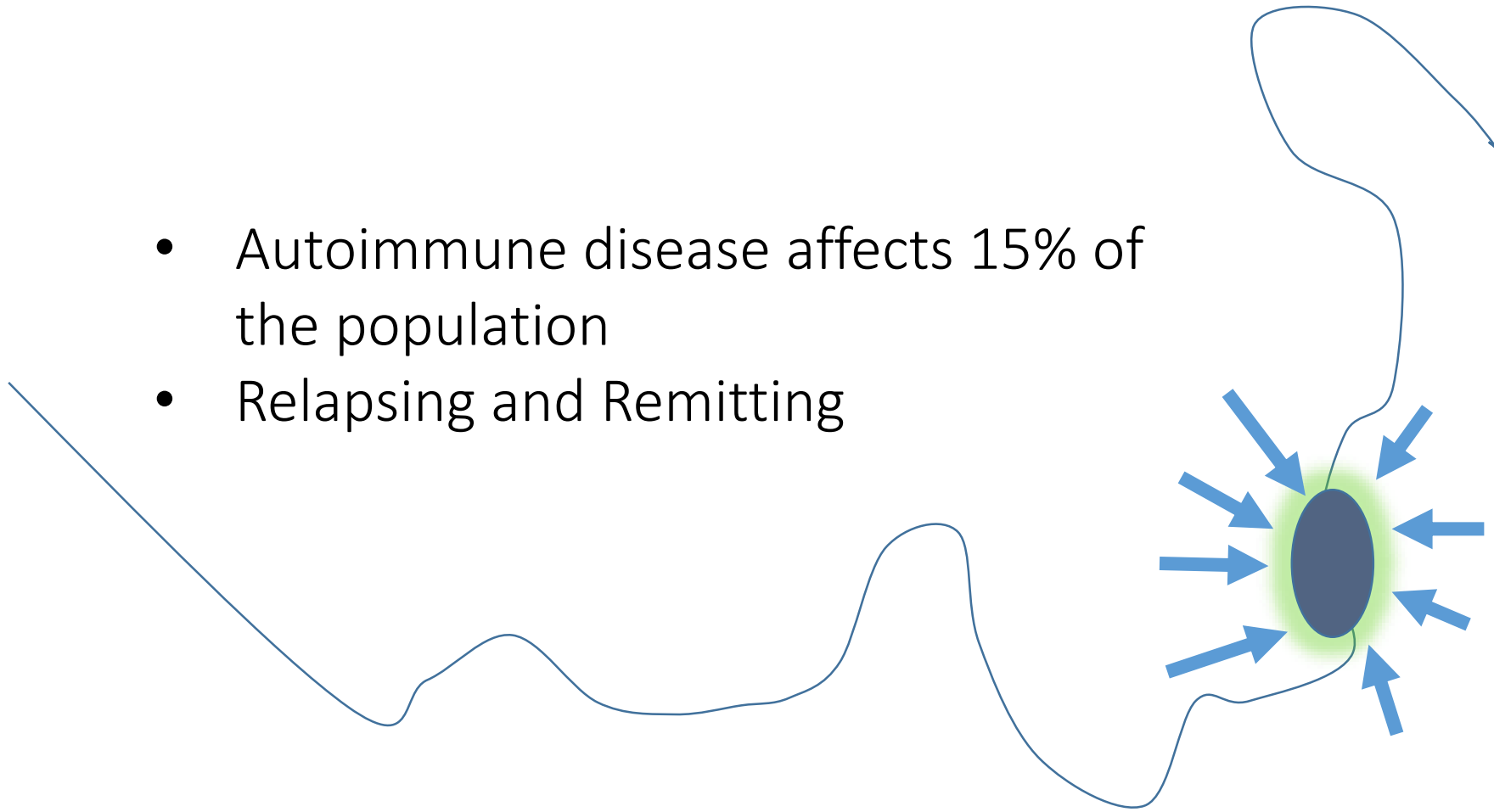# A big data approach for understanding environmental risk factors for autoimmune disease flare

Brian Reddy

Trinity Health Kidney Centre / ADAPT

- Autoimmune disease affects 15% of the population
- Relapsing and Remitting

Cloudy with a chance of pain!

- The world's first smartphone-based study to investigate the association between pain and the weather
- Anyone in the UK with arthritis or chronic pain and aged over 17 can take part
- The information could be used for generating pain forecasts, allowing people to plan their weekly activities

It's a mystery that's perplexed people for over 2,000 years, but now University of Manchester scientists are on the verge of working out if the weather affects pain in people with arthritis and other conditions, all thanks to the British public and their smartphones.

Cloudy with a Chance of Pain, which launches today (26 January) is the world's first smartphone-based study to investigate the association between pain and the weather. The study will be carried out during 2016 using a smartphone platform called uMotif which people will use to record how they're feeling, whilst local weather data is automatically collected using the phone's GPS.

Anyone in the UK with arthritis or chronic pain and aged over 17 can take part. All participants need is a smartphone.

Click here to download the app and take part.

Dr Will Dixon, Director of The University of Manchester's Arthritis Research UK Centre for Epidemiology and Honorary Consultant Rheumatologist at Salford Royal NHS Foundation Trust, came up with the idea. He said: "This question has been around for more than 2,000 years, but it's only now with widespread modern technology that we have the ability to answer it.
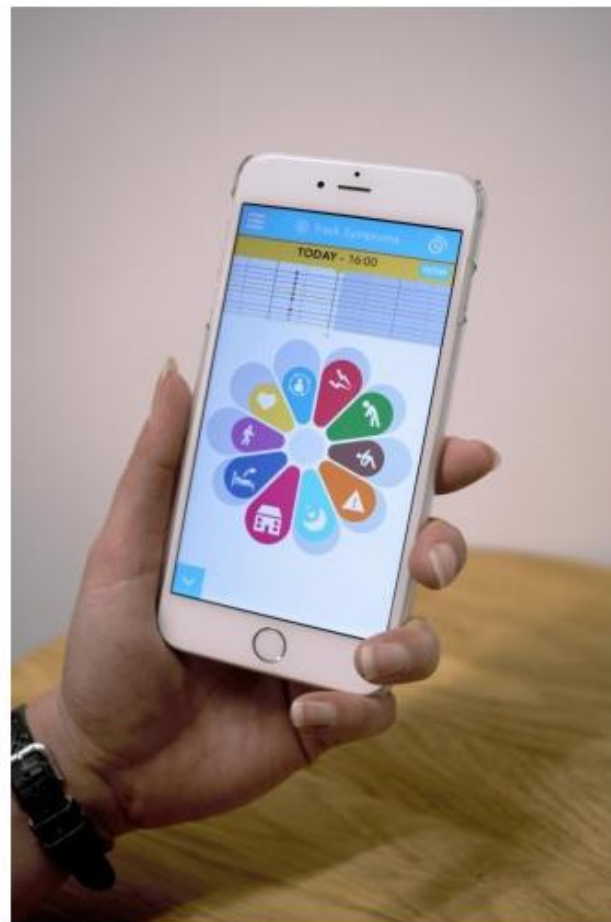
Background

- ANCA vasculitis is a relapsing and remitting, rare autoimmune disease.

- 'Flares' of the disease can result in rapid kidney impairment and destruction of other organs.

  - These cannot currently be predicted with accuracy.

- Treatments to reduce the risk of flares occurring are toxic, and ideally should not be taken unless necessary.

- Epidemiological data support a strong environmental impact relating to flare risk.

  - But it has proven difficult to identify exactly what environmental components are relevant, partially due to the rareness of the disease.

- The **AVERT** study is aimed at stratification of the risk associated with disease at the patient level.

The role of the exposome

- There have been major advances in the last 20 years in understanding human health, such as mapping the human genome.
  - But this has thrown up many new problems!
- Genetics now thought to be responsible for as little as 10% of disease.
- Environmental causes are believed to cause 70-90% of disease (Juarez et al. 2014)
- The exposome the measure of all the exposures of an individual in a lifetime and how those exposures relate to health.
- Application of IT infrastructure to environmental monitoring data could facilitate major advances in patient management.
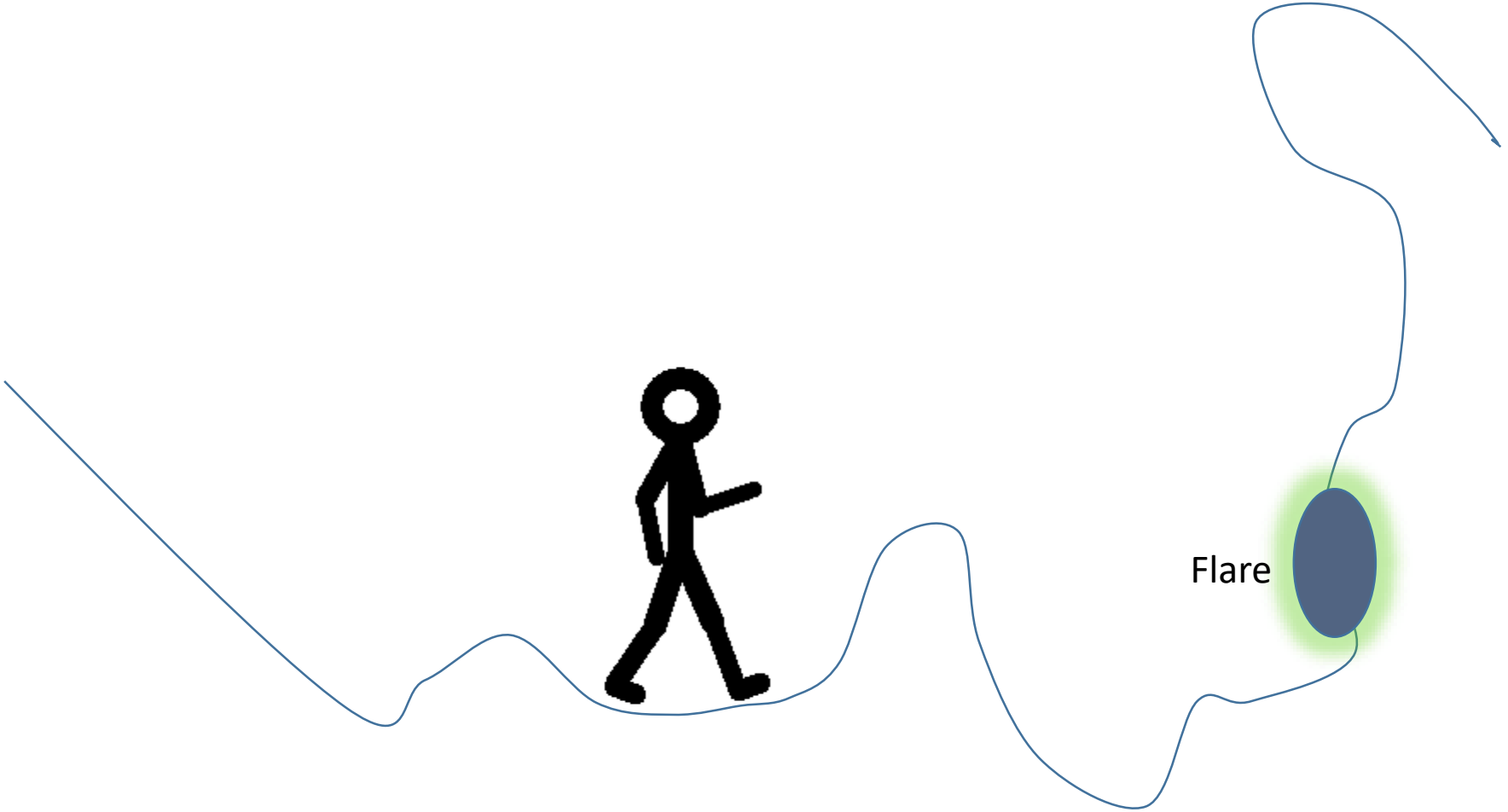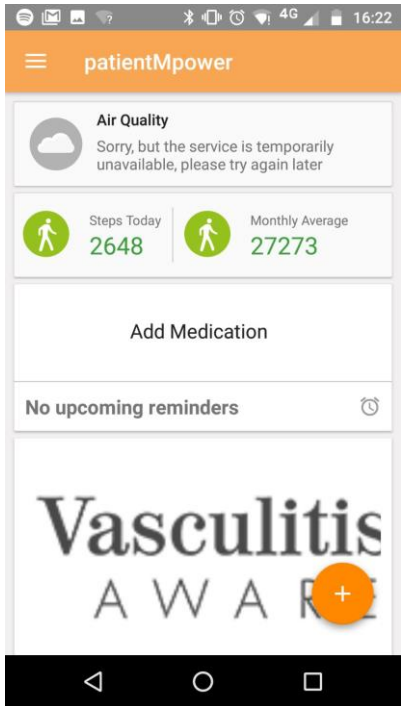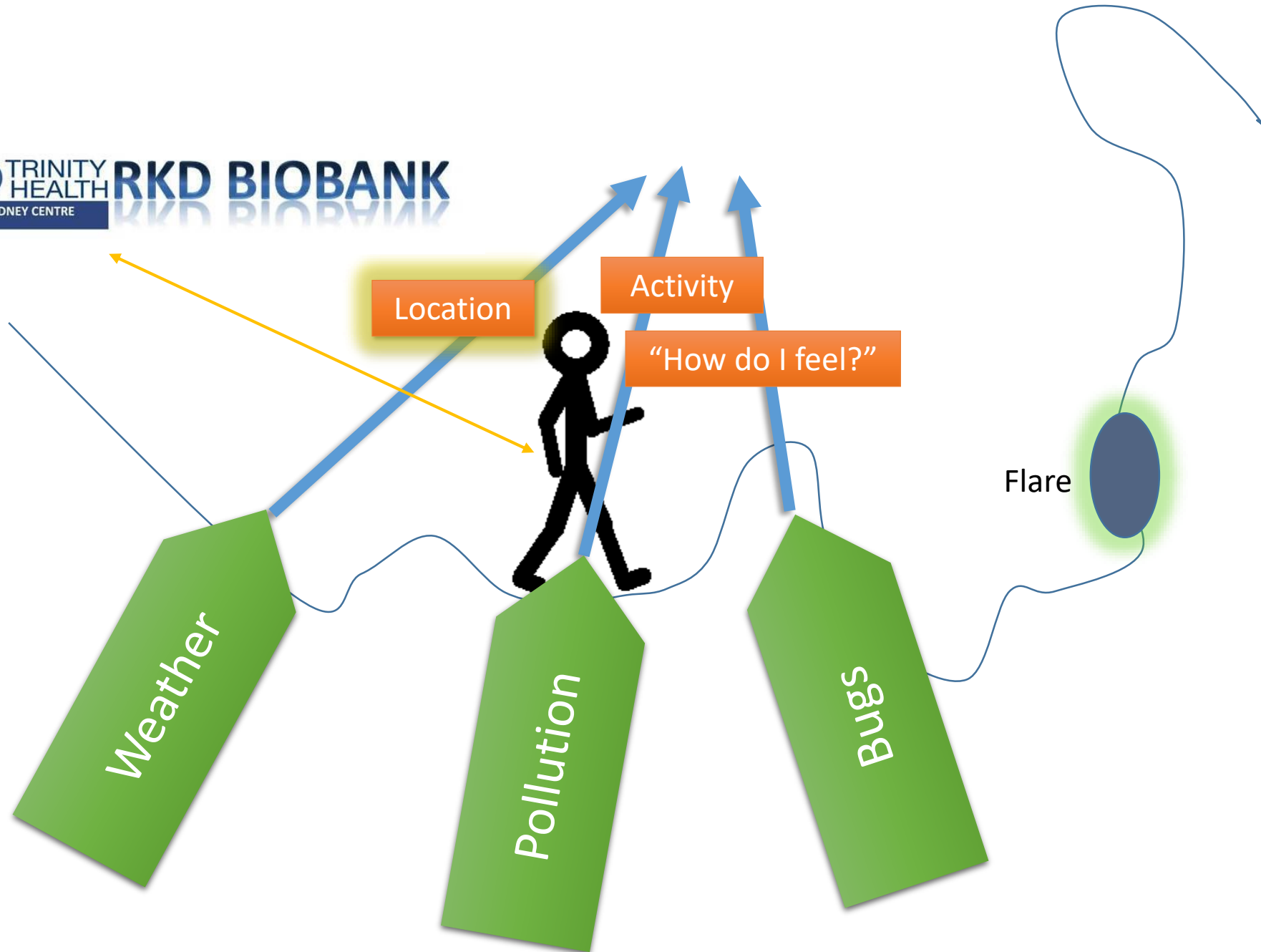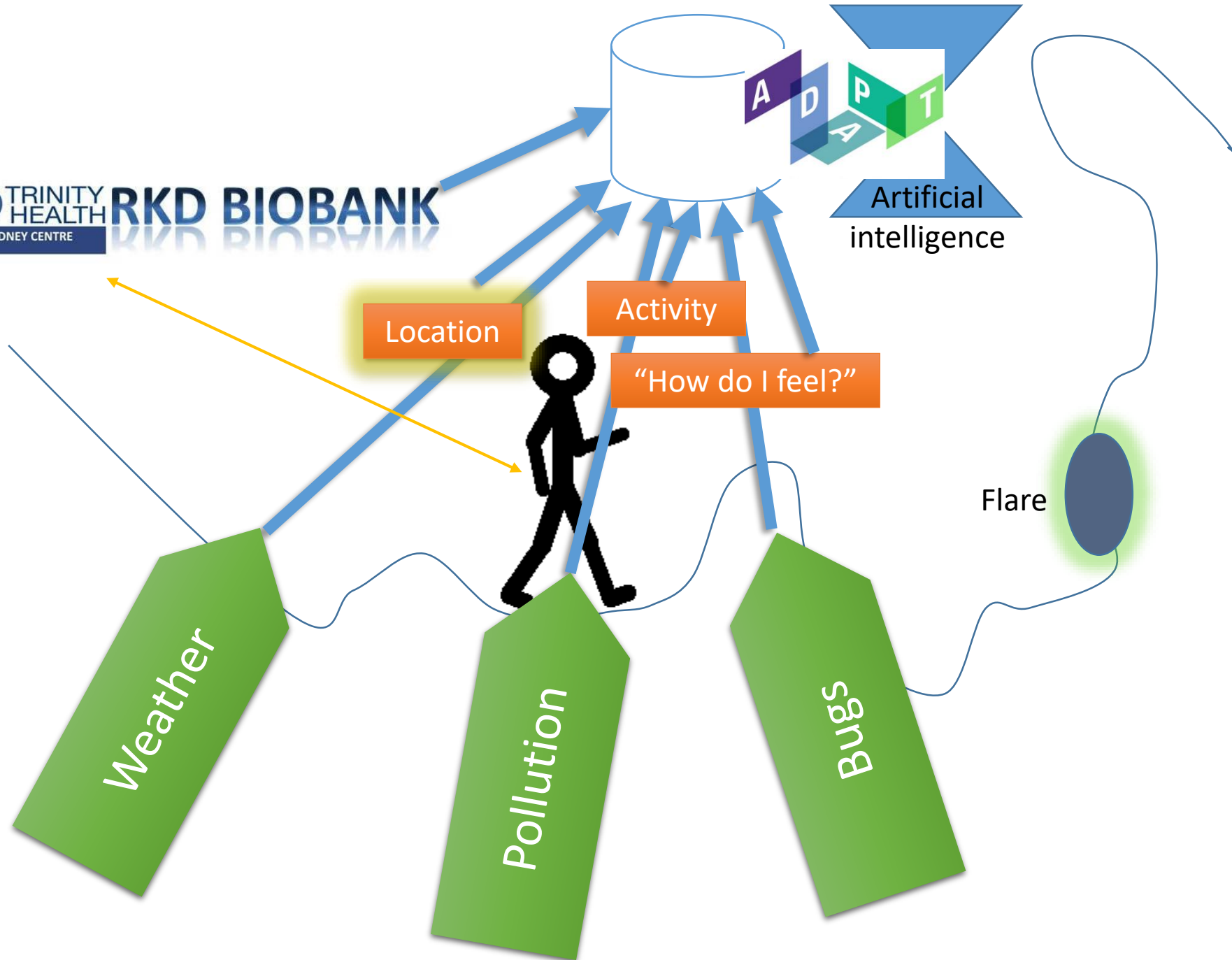
Flare
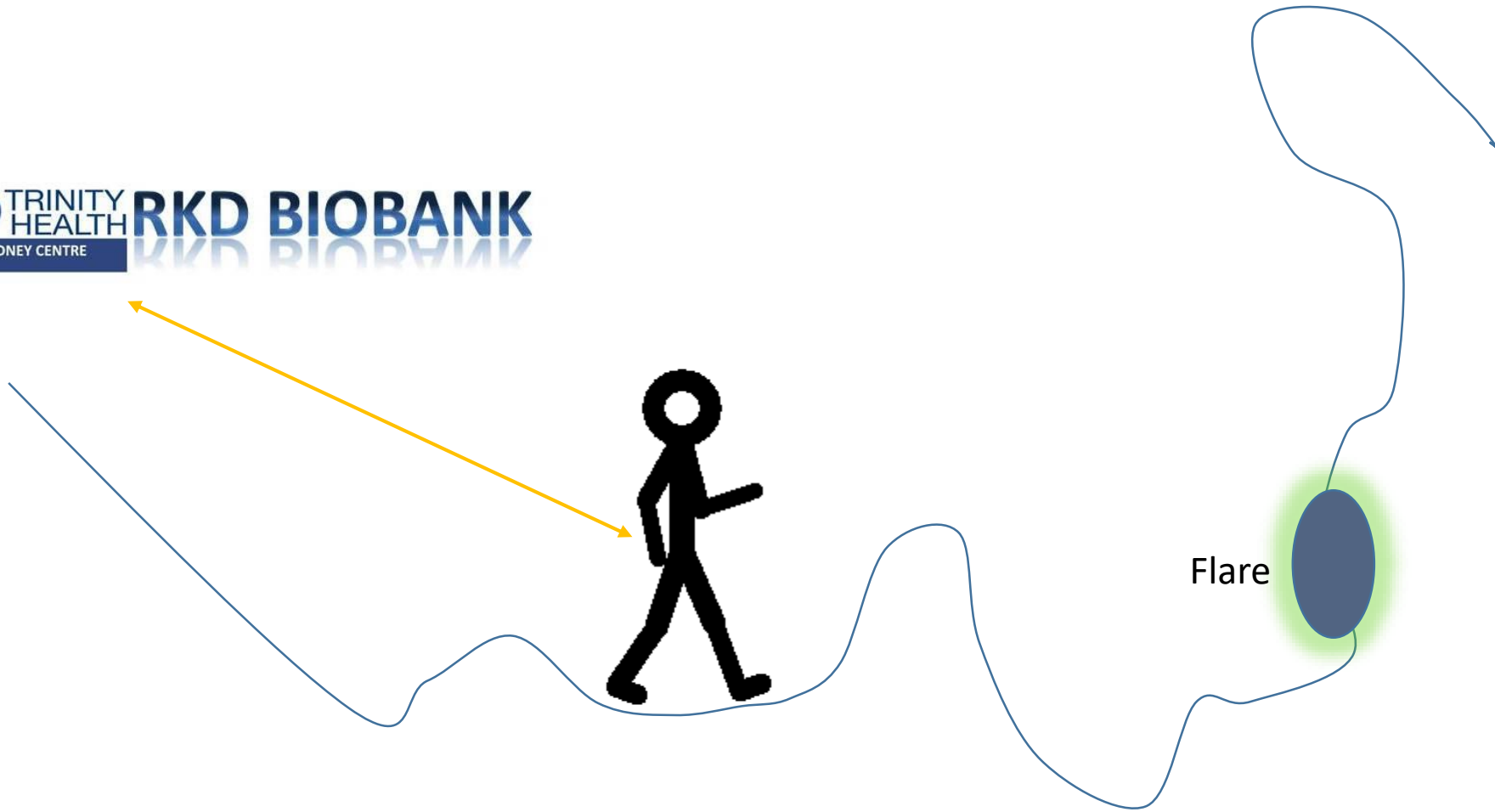
Flare
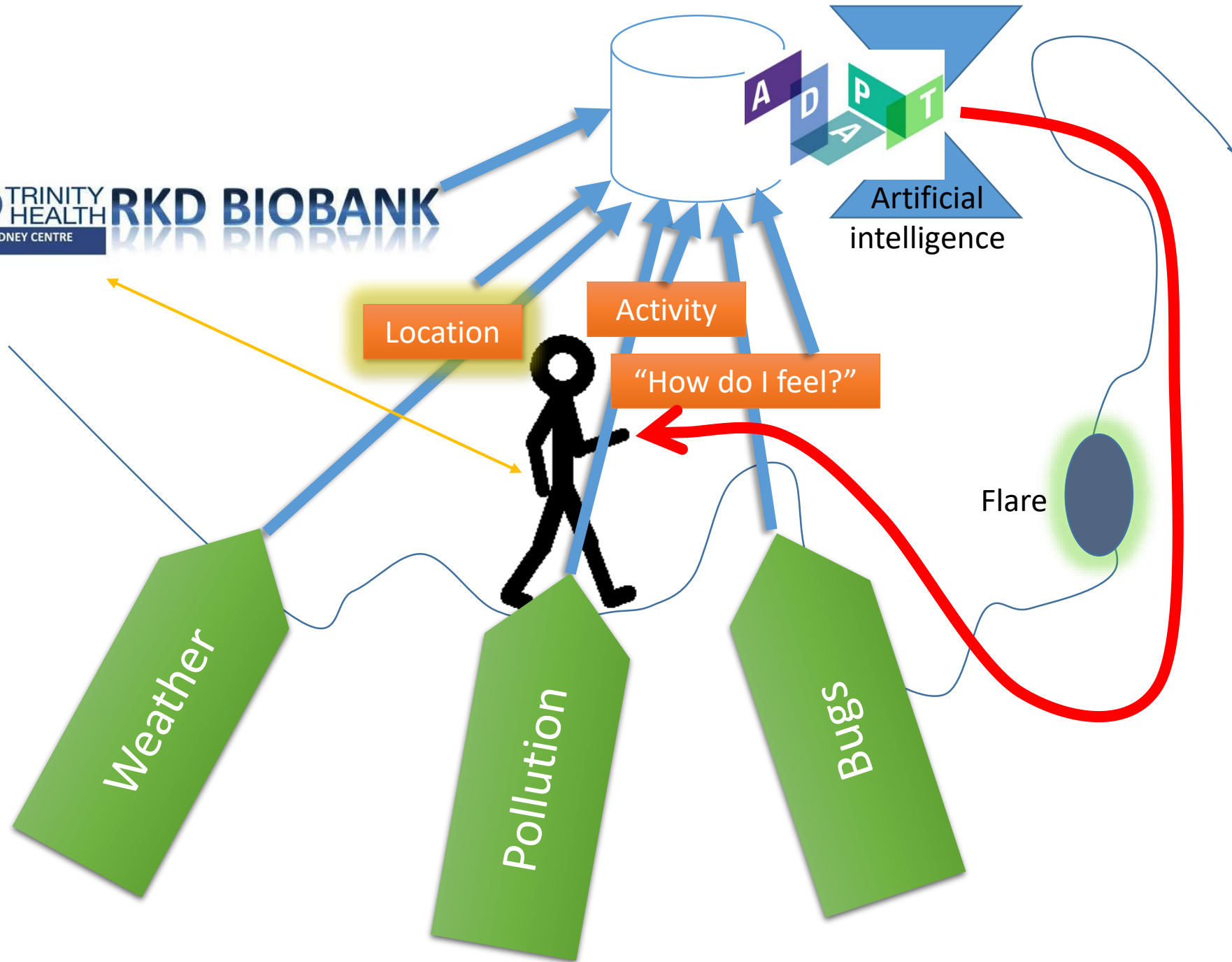
Study hypothesis

1. Interaction between patient-level and environmental factors leads to relapse of PR3 ANCA vasculitis;

2. That the biological signal associated with this can be resolved by integrating highly granular longitudinal data using non-parametric regression, and;

3. These algorithms will be sufficiently robust to inform predictive machine learning at individual patient level.

Semantic web approaches

- There are specific challenges associated with analysis of data at this scale, with careful planning required around the management and storage of data.

- We are using a linked data model, the Resource Description Framework (RDF), to model and manage this information, which allows:
  - for different data sources and files to be easily understood (and combined) by both humans and machines;
  - quicker and more intuitive querying of data;
  - For the capturing of geographical features, and;
  - semantic reasoning to be applied to generate further enriched data.

The AVERT vision

- Sharing

- Role of RDF

- Ethical issues
    - "Impossibility of anonymity"

Analysis

- At most half of patients in study will have flares each year

- This is a classic "large p, small n" study. It will take some years to build up sufficient numbers of flares for definitive answers.

- There are virtually limitless variables available in principle that could include. Already it is clear that complexity will arise when trying to take into account:
    - Relevant/plausible lag times for explanatory variables
    - Interactions between explanatory variables

- Our initial plan is to employ Bayesian Additive Regression Trees to find patterns, but will have plenty of time to plan.

- For the time being this remains very much a project focussing on building up the infrastructure to allow this analysis in future, though we are testing similar approaches on historical data currently.

Conclusions

- Flares of ANCA vasculitis - like many autoimmune diseases - appear to be associated in time and space with poorly understood environmental factors.

- The community of vasculitis patients in Ireland is an engaged group of an appropriate size to attempt to gather sufficient data to better understand these.

- The ultimate goal of AVERT is to allow patient-level risk factors to be better identified, allowing benefits and risks of anti-flare treatments to be weighed up.

- More broadly, the approach may represent a new paradigm in managing chronic conditions governed by interaction between patient-level factors and their environment.

- It may also blaze the trail for application of similar approaches for other (more common) autoimmune diseases

Thanks!

Chronic Disease Informatics Group

Introduction

The Chronic Disease Informatics Group (CDIG) is interested in the application of novel statistical approaches to better understand the impact of environmental influences on chronic disease. The group brings scientists from the disciplines of medicine, statistics and computer science together.

In particular, the group is concerned with better stratifying the patient-level risk associated with flare ups of ANCA-associated vasculitis. This condition can cause multi-organ failure as a consequence of overwhelming necrotising inflammation affecting small blood vessels.

Particular thanks to:
- Mark Little
- Lucy Hederman
- Brett Houlding
- Alan Meehan
- Jason Wyse
- Declan O'Sullivan

Contact:
reddybr@tcd.ie
@imirtpele