SAMPLE

Building Tomorrow's Researchers

The Protocol's Keeper in the Machine Era

Aiman JoeHanz

Free sample: Introduction & Chapter 1

putrabooks.github.io

Buy the full book:

Carousell carousell.sg/p/building-tomorrow-s-researchers-ebook-pdf-1471387822

Ko-fi

Payhip

Contents

Introduction

Chapter 1 — The Protocol

Chapter 2 — The Two Folklores

Chapter 3 — The Way In

Chapter 4 — The Working Day

Chapter 5 — The Craft

Chapter 6 — The Churn

Chapter 7 — The Conversation

Chapter 8 — The Instruments

Chapter 9 — The Year

Chapter 10 — The Long Arc

Introduction

A research lab holds considerably less than the public imagines, since what a visitor is shown is a bench with a fume hood, a freezer that alarms when the power fails, a cabinet of labelled tubes, four screens, a whiteboard of half-erased arrows, and a notebook that nobody outside the room will ever read. The building's other floors carry the myth, which runs through the corridor where a genius paces until the idea arrives, the seminar room where the breakthrough is announced, and the press office that turns a coefficient into a headline, while the bench is where the profession actually happens and where its most consequential object is also the least photogenic thing on it, a document written before any data existed that says what will be measured, how, on whom, and what result would count as a failure.

This volume follows the researchers, which means the postdoctoral scientists, the clinical trialists, the field epidemiologists, the laboratory and social-science investigators, and the analysts who work on publicly funded studies that nobody will ever interview them about, and its claim, argued across ten chapters, is that a single artefact keeps the whole trade honest without being the paper. It is the protocol, the plan committed to before the answer is known, which fixes the question, the measure, the analysis, the stopping rule and the date, and protocol first means the study can fail, and a study that can fail is the only kind that can teach anything. Everything the trade fears, from the unrepeatable result to the finding that dissolves when someone else looks to the conclusion that arrived before the data, turns out to be a protocol problem wearing a statistical costume.

What the Volume Promises

Three things are promised here: working researchers get their trade named precisely, including the long stretches that are mostly waiting; people at the door get the routes, the timings and the numbers the job market actually pays; and every reader gets a usable test for the science they meet daily, since the public is served a result a day and owns no protocol for reading it. The reader who finishes this volume will be able to open almost any study and ask what was written down before the data arrived, which is the question that separates evidence from advertisement.

Who This Volume Serves

Four groups are served: practising researchers who recognise their own corridor in these pages, students and technicians weighing a life in the lab who deserve the arithmetic before they pay an application fee, the clinicians and

teachers and engineers and civil servants who use research and are routinely handed its conclusions without its conditions, and the reader with no trade connection at all, who has met a hundred headlines opening with the words a new study and deserves to know what such a study usually looked like.

A Note on the Volume's Cases

Every case in the volume is constructed, since the lab of the ninth chapter, its staff, its funder and its year are drawn composites assembled from patterns common in the trade, and the studies that go unnamed in earlier chapters were built to carry their lesson rather than borrowed from anybody's file, for three reasons stated plainly: confidentiality, because unpublished data belongs to the people who consented to provide it; teaching, because a drawn study can be simplified without losing its shape; and self-audit, because drawn cases can be kept consistent, which is why the statistician of the fifth chapter reappears in the ninth and why the arithmetic is meant to agree. One limit stays honest: no invented protocol conveys what it costs a real investigator to write one when the career depends on the answer coming out a particular way.

The Volume's Place in the Collection

Every volume in this series examines one discipline, the keeping of dated promises, as a different trade practises it under pressure from a different machine. Readers arriving from an earlier volume will recognise the furniture: an artefact that holds the trade together, two folklores, a working day, a churn, a constructed year and a long arc. Research gives the discipline its purest test, since in this trade the promise is made to strangers, in advance and in writing, and its keeping is checked by people who never met the ones who made it, which is all that peer review amounts to, a promise checked by the uninterested, and the trade's whole dignity, together with most of its difficulty, follows from that arrangement.

One sentence travels with the reader through the volume: examined work endures, and in this trade the examination that matters occurs only after you have committed, beforehand and on paper, to what you expected.

How the Volume Argues

The argument is cumulative rather than declarative, and the reader deserves to know the method before it is used: each chapter takes one part of the trade, describes what it looks like from the inside, names the failure that recurs there, and shows a repair that costs less than the failure, and nothing in

the book depends on a scandal, because scandals are rare while the ordinary erosion of standards is not. This book addresses the researcher trying to do sound work inside an imperfect system, not the reader in search of a villain.

The Machine's Share

The era's instruments appear in every chapter and in none of them as a saviour, and the introduction owes the reader the distinction before the argument gets far, which is that a model can draft a protocol section, search a literature, propose an analysis, generate a figure and write a fluent paragraph about a result it has never seen, and that none of those capacities touches the thing this book calls the work. The promise is still made by a person to a question, the protocol is still dated by somebody who will be answerable for it, and the anomalous value in a spreadsheet remains evidence of something a human being has to look at twice.

What the machine changes is the price of the parts around the promise: the literature review that took a month takes a week, the code that took a fortnight arrives in an afternoon, and the first draft of the discussion writes itself. The chapters will show what happens to that reclaimed time in honest and dishonest hands, since the trade's oldest failure mode is not laziness but the substitution of an easy artefact for a hard judgement, and the machine makes easy artefacts cheaper by the day.

Chapter 1 — The Protocol

The work of research begins with a document that has an unusual property: it can be wrong. A plan for a study, written down before the measurements exist, commits the investigator to a question, a population, a comparison, a duration, a measurement method and a decision rule, and the last of these is the one that gives the trade its spine. A study whose analysis can be adjusted once the data arrive is not a test of a belief; it is a search for a presentable version of one.

What the Work Actually Is

Ask a researcher what they did last week, in order, and the myth collapses under the answer, and there was a grant report due Friday. There were two failed runs of an assay whose reagent lot had changed, and there was a long meeting about whether the control group was really comparable, and a phone call to a statistician that resolved it. There was one hour with a specimen, three hours of code, and four hours reading papers that made the question look worse rather than better. Somewhere in that week was the activity the myth prefers, thinking, which took place mostly in the shower and during the walk back from the car park.

The trade's honest self-description is the manufacture of trustworthy claims under conditions of uncertainty and a budget, and it is manufacturing, in the sense that it has process control, tolerances and waste, and it is not manufacturing in the sense that the product is a sentence the investigator does not yet believe.

The Keeper

The protocol is the trade's keeper, and the argument for it is structural rather than moral. A researcher under pressure has an interest in the outcome, so so does the funder, the institution, the journal and the press office, and none of them is wicked for it. When an interested party examines data without a prior commitment, the flexible places get used, usually without conscious dishonesty: the outlier is excluded for a considered reason, the outcome is chosen from four because one is more clinically sensible, the subgroup becomes the headline because it turned up a signal. The protocol exists so that those decisions were made before the data could argue.

This volume calls the consequence the pre-registration principle, though the trade has other names for it, and it says a study is a promise to a question, and the promise is credible only if its terms are dated before the answer.

The Four Breaks

Across every research setting, the same four failures recur, and each will get its own case in the first half of this book.

The first is the unregistered hypothesis, where the plan is quietly rewritten after a look at the data, and the study keeps its title and loses its test. Nobody usually plans to do this; they plan to be flexible, and flexibility is what the break is made of.

The second is the turned page, the result that nobody replays. A finding is published, the sample is argued about, and no one runs the study again with fresh subjects, because a replication is a career cost without a novel claim, and the literature grows in breadth and stays thin in depth.

The third is the story's pull, the mechanic that converts a careful estimate into a shape with a hero, a villain and a lesson, and it happens in the abstract, in the press release and at the kitchen table, and it happens because narrative is how people remember, not because anyone lies.

The fourth is the hunt for significance, the flattering measure of this trade. An effect small enough to be explained by noise is hunted through analysis choices until a threshold is crossed, and the crossed threshold is then reported rather than the journey.

The Cases in Brief

A postdoctoral fellow runs a promising experiment and the result is null, and she looks again with a different outcome measure and finds something. Her supervisor is sympathetic, the revised measure has a defensible rationale, and the published paper does not say which came first. Later, two other labs fail to find the effect, and the break here is not fraud; it is an unregistered hypothesis, and it consumed four years. The repair is the dated analysis plan, filed before the first subject was enrolled.

A field study reports a large association, and no one replicates it, because a replication earns no credit and consumes a season, and after eight years a policy is built on a single site. The turned page is the break, and the repair is the trade's least glamorous institution: a funder line item for confirmation, paid for like the equipment.

A trial finds a modest benefit, and the university's release says the trial found a breakthrough, adds a patient story, and the clinic's phones ring for a year with people asking for a treatment that was tested on 214 people over twelve

weeks. The story's pull did the damage, not the trial, and the repair is a release drafted by the authors and cleared before the conference.

A doctoral student needs a chapter to finish, and three outcomes were measured, one reaches a threshold, and the thesis reports that one. Nothing in the write-up is false and the effect is probably real, and the hunt for significance still happened, and the field inherits one overestimated number that will be cited for a decade.

Why This Volume Begins Here

Because everything else in this trade is downstream of it. Career structures, funding models, journals, prizes, press offices and now the models that draft manuscripts all press on the same point, and the protocol is the point where the trade decides whether it is producing knowledge or producing publication, and the remaining chapters describe the profession that lives with that decision daily.

What the Keeper Is Not

It is not a form, though many institutions have made it one. A fifty-page template completed at the last minute from memory of what the study intends keeps nothing, and the test of a protocol is a sentence in it that the investigator would rather not have written, usually the one naming what result would mean the idea is wrong. A study that cannot disappoint its author is a demonstration, and demonstrations belong in sales.

It is also not a guarantee. A well-planned study can still be wrong, and often is, because reality is bigger than any plan, and the claim being defended is narrower and firmer: a dated plan makes an error findable, and an error nobody can find is not a temporary setback but a permanent contribution to the literature.

The Interim Look

Nearly every long study takes an interim look, and the look is where a well-planned trial can quietly become an unplanned one. A monitoring committee sees the accumulating data before the investigators do, precisely so that the accumulating data cannot steer the investigators, and the arrangement's logic is worth stating because it is the protocol principle applied to temptation: the person who will have to live with the result does not get early sight of it.

The rules that keep the look clean are few, and the number of looks is fixed in

advance. The threshold for stopping is set before the first subject enrols, and the committee is independent, and its recommendation is recorded even when the investigators disagree. Without those three, the interim becomes a series of peeks, and a study that could have stopped honestly instead wobbles toward whatever the last look suggested.

The repair is unromantic: put the stopping rule in writing, name the committee, and accept that a study may end early because the answer arrived sooner than expected. That sentence is the whole discipline of the interim look, and it costs an hour to write and a decade to learn.

The Audit Trail

Every claim this chapter has made about the protocol assumes something humbler than integrity: a trace. An audit trail is the dated sequence of who changed what, when and why, and its absence is the difference between a study that can be examined and a study that must be believed.

The trail is made of small things, and the protocol's version numbers. The amendment log with its reasons, and the informed-consent form's versions and the date each came into use. The analyst's commits, and the minutes of the meeting where the primary outcome was defended. None of these is impressive on its own, and together they answer the only question an auditor truly asks, which is whether the study's story could have been told differently on any given Tuesday.

A lab without a trail is not necessarily dishonest, and it is uninsured, and the insurance is what makes a career's work survivable when someone later doubts it. The chapters that follow will show the trail catching errors in ordinary weeks; here the point is only that the trail must exist before it is needed, since a trail assembled afterwards is a reconstruction, and reconstructions are what the trade's four breaks are made of.

The Sample Size's Logic

The most misunderstood number in a study is the one in its recruitment target, and the misunderstanding is worth a section because it produces both waste and false confidence. A sample size is not a measure of ambition, and it is the output of an arithmetic that starts with the smallest difference worth caring about and the amount of noise the measurement carries.

Four inputs decide it, and the effect size that would change practice, chosen before the study begins and defended in the protocol. The variability of the

measurement, estimated from prior work rather than invented, and the tolerable chance of a false positive, and the tolerable chance of missing a real effect, both stated as numbers. Change any of them and the target moves, which is why a study that quietly triples its effect assumption to fit its budget has changed its question rather than its plan.

The chapter's practical warning concerns the small study most of all. A sample too small to detect a modest effect will usually return an interval spanning everything from harm to benefit, and the temptation at that point is to report the point estimate and neglect the width, and the honest report prints the whole range and says that the study was underpowered, which is a finding about the design rather than about the world, and the field needs it stated rather than buried.

The Pre-Specified Subgroup

Subgroup analysis is where the trade's good intentions meet its arithmetic, and the honest protocol handles it in advance rather than after the fact. A subgroup is a question about whether an effect differs across types of people, and the mathematics of testing many such questions guarantees that some will look interesting by chance alone.

Three rules make subgroup work usable. Name the subgroups in the protocol, and keep the list short enough that each is justified by a mechanism rather than by curiosity. Report the interaction rather than the separate results, since an effect in one group and none in another is compatible with no difference at all until the interaction is tested, and label everything else exploratory, which is a respectable category when it is honest.

The break this rule prevents is the discovery announced at a conference: a treatment that works in women under fifty, found by looking at eight slices of a trial designed to answer none of them. The repair is not to stop looking, since subgroups are where clinical nuance lives, but to look in a way that would survive being asked, before the event, what result would have convinced you.

The Registry's Fields

A registry entry is a short document and every field in it is doing work, which is why the trade should read one properly at least once, and the title names the question rather than the hoped-for answer, and an entry whose title promises benefit is already advertising. The dates say when the plan was fixed and when the first participant arrived, and the gap between them is the honest

measure of whether the plan preceded the data.

The outcomes field is the decisive one. Primary comes first, singular, and defined; secondary follows. When a published paper promotes a secondary outcome to the headline, the registry is the document that shows it, and the reader who checks is doing the trade's most useful act of citizenship.

The rest repays attention too, and the sample size declared against the analysis intended. The arms described, and the sponsor named. The contact details of a person who will answer an email. A field left blank is usually a decision deferred, and a deferred decision is one that the data will later make, which is exactly what the plan existed to prevent.

The Question's Size

A research question can be too large in a way that is invisible until year three, and the discipline of sizing it is the first craft a newcomer should learn. A question that cannot be answered by a person with the resources on hand is not ambitious; it is unstarted.

The sizing test has three parts. Can the outcome be measured with an instrument the team already has access to. Can the population be reached within the recruitment window, and can the question be answered in the time the funding provides, allowing for everything that has historically gone wrong. A question that fails any of the three should be cut in half rather than defended, and the cut version is usually better science, because a smaller question gets a cleaner answer.

The trade's habit of praising large questions is a folklore inheritance, and the practitioners who produce the most durable work are the ones who spent their careers asking questions small enough to finish, and the footnote in the field's history that everyone remembers is a small question that was answered, and the ones nobody cites are the grand ones that were never closed.

What a Failure Looks Like

The protocol's hardest sentence is the one that describes the result that would end the idea, and the trade's training spends almost no time on how to write it, and it is not a formality. A study whose author cannot describe a refuting observation is a study that will interpret whatever arrives as support, and the interpretation will be honest and wrong.

The sentence has a shape, and if the difference between groups is smaller

than X, or the interval includes no meaningful benefit, or the measure does not move in the pre-specified direction at the pre-specified time, then the hypothesis is not supported and the intervention should not proceed. Written out, the sentence is uncomfortable, which is the test of whether it is real.

The chapter's observation is that teams who write it find the exercise clarifying rather than demoralising. A described failure is a boundary, and boundaries make the work finite, and the alternative, an idea that no possible result could refute, is not a scientific position at all, and the trade's more mature practitioners have learned to be suspicious of their own favourite hypotheses for exactly this reason.

The Team's Protocol

Protocols are written by committees and committees have a known failure mode: the document becomes the union of everyone's concerns and the analysis is nobody's decision. A protocol with nine authors and four primary outcomes is not thorough; it is unresolved.

Three practices prevent it, and one author owns the document and has the final pen, which is a role rather than a hierarchy. The analysis section is written by whoever will run it, with the statistician's line-by-line agreement recorded, and the contentious choices are minuted, with the rejected option and the reason, so that the decision survives the meeting and can be defended later when someone asks why the study did not measure the obvious thing.

The minuting matters more than it sounds, and the trade's disputes about a study usually turn out to be disputes about a decision that was made in a room and never written down, and the group that keeps the record of its own arguments can settle a reviewer's objection in a paragraph rather than a correspondence of six months.

The Secondary Question

Every study collects more than its headline measure, and the extra measures are where the protocol's resolve is tested in the quietest possible way, and the plan lists the primary outcome and a set of secondary ones, and the temptation that arrives with the data is to let whichever measure moved best become the story. Practitioners have a phrase for what the result is called, outcome switching, and they have a counter, which is the same defence every chapter of this volume returns to: write the hierarchy down before the first participant is enrolled, number the outcomes, and let the numbering survive

the analysis.

The worked version is dull and takes ten minutes. A trial of a walking programme lists six outcomes, with the six-minute walk first, and the three-month data arrive and the six-minute walk moves by less than the trial was powered to detect, while a quality-of-life score moves nicely. The undisciplined paper writes itself in an afternoon: the programme improved quality of life, with a note that the walking distance was not significantly different. The disciplined paper says the primary outcome showed no detectable effect and reports the secondary finding as exploratory, labelled as such in the abstract, and the difference between those two papers is not honesty as a virtue but a numbered list written in the spring.

The practice has a corollary for readers of research, since most readers in the trade are also consumers of other people's studies, and when a paper's abstract leads with a secondary measure and buries the primary result in the fourth paragraph, the reader is entitled to ask who decided the hierarchy and when, and the registry entry answers the question in one line if it was answered at all. A field where readers check the registration before trusting the abstract is a field where the abstract starts telling the truth, and nothing in the checks requires the reader to be a statistician.

The Stopping Rule

A study that can end early needs a rule for ending, written before anyone knows what the data will say, and the rule is the protocol's most exposed clause, and the interim numbers arrive at a committee meeting in a sealed envelope or a locked file, somebody does the arithmetic, and the study either continues or halts. The honesty of the halt depends entirely on whether the threshold was set when the investigators were still ignorant, because a threshold chosen after the data have been seen is not a rule at all, and it is a decision wearing the costume of one, and the difference is invisible in the published paper, which will simply report that the trial stopped early for benefit.

The trade has learned this the expensive way. Trials stopped early on a favourable signal are notorious for overstating the effect, since the stopping point is often the high-water mark of a noisy measurement, and the honest protocol therefore names the person who will look, the number of looks, the threshold at each, and the consequence of crossing it. The rule's second function is less discussed and just as important: it tells the disappointed

investigator, in advance, that a halt for futility is a result rather than a failure, and it gives the team a sentence to say to the funder that does not depend on improvisation.

AUTHOR'S NOTE

This book describes the daily work of research, and it was written on the assumption that a reader deserves to know what a trade actually does before admiring, joining or reforming it. The protocol is the trade's keeper in these pages: the dated plan committed to a question before the answer is known, written by people who will be answerable for it, and the honesty of that plan is treated as the discipline's whole grammar, learned in a first study, priced in working weeks and audited one pre-specified endpoint at a time.

Every worked case in this volume is constructed. The labs, trials, committees, institutions and people are drawn from the ordinary weather of the trade rather than from any individual's record, and their numbers are illustrative rather than measured. Where the book describes a method, it describes the method as practitioners use it; where it describes a failure, the failure is one the trade already knows by name.

The machines appear in every chapter and are never asked to carry the argument. A model can draft, search, propose and write. It cannot make a promise to a question, and it cannot be held to one, which is the distinction this book returns to whenever the era's enthusiasm arrives early.

The author thanks the researchers, trialists, statisticians and laboratory staff whose ordinary conversations shaped the volume's sense of what the work costs, and any error of fact or emphasis remains the author's own.

ABOUT THE AUTHOR

Aiman JoeHanz writes the Building Tomorrow's People series. The final volume in the current run belongs to the researcher, the trade that keeps the protocol: the plan committed to a question before the answer is known. This book follows the grant, the pre-registration, the analysis that was specified and the one that was not, the reviewer who asks for a different study, the model that now drafts the section, and the long, unglamorous discipline of writing down what you intended before you find out what happened.