coordinator@medconfidential.org

9 August 2026

Dear Dr Myles,

medConfidential has long argued that HDRUK and UK Biobank share a culture, the consequences of which for UK Biobank are now documented in Hansard. You'll also be aware of the re-identification of Biobank participants which should put to rest any question about whether the data you provide remains identifiable and the need for controlled access.

We are concerned that CPRD has imitated UK Biobank's flawed approach to governance and access control in your partnership with HDRUK. CPRD has delegated Safe People and Safe Projects decisions to an entity with no obligation to keep CPRD's promises and with no capacity to do so.

CPRD claims "Five Safes" adherence while seemingly ignoring them in practice. As with UK Biobank, HDRUK has redefined the Five Safes to permit the reckless choices already made.

None of the below should be a surprise to you. But can you confirm whether CPRD is happy that the reassurances you have on your website for practices and patients remain accurate? Are all your recent choices entirely compliant with the legal agreements under which GPs gave you their patients' data?

CPRD's assurances may be undeliverable

The CPRD website says:

- "Researchers must submit a study application to access anonymised patient-level CPRD data" (here)
- "Before they are able to access anonymised CPRD data for approved research purposes, all research organisations must demonstrate, to CPRD's satisfaction, that they handle data in line with the standards required by UK/EU/EEA data protection legislation, and that their information security measures align with established data security standards" (here)
- "Anonymised CPRD data can only be accessed and used by approved researchers for research to benefit patient and public health." (here)

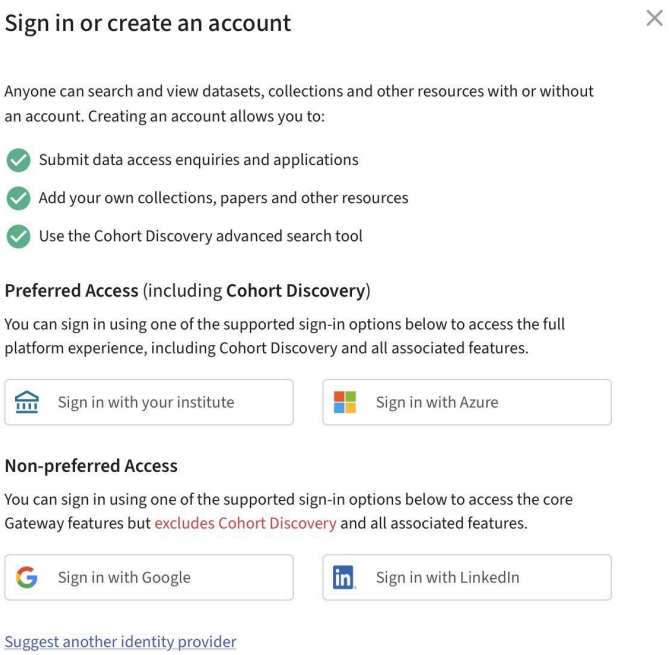
Because HDRUK quotes you in their announcement, you'll hopefully already be aware that the HDRUK website has a new feature on their gateway: "The federated onboarding of anonymised CPRD records for over 49 million patients into Cohort Discovery means that researchers can now identify larger, more representative study cohorts drawn from primary care data."

The gateway “cohort discovery” feature does not appear to satisfy your promises to patients and GP data controllers from whom you received data. Without publication of your agreements with NHS England, there are no public sources to confirm whether you have made HES or other NHS England datasets available for processing in this way, and whether your agreements with NHS England would permit you to do so (NHS England staff interested in this paragraph may wish to pay unusual attention to the postscript question to this letter).

We note the assertion from HDRUK that “The Cohort Discovery Service follows the Five Safes Framework”, a claim that sounds like the assertion UK Biobank made while their supposedly “safe” environment had a ‘download all the data’ function (that UK Biobank leadership knew about).

The gateway guidance states that it is HDRUK, not CPRD, which determines who is “a recognised researcher, NHS analyst or equivalent” and “Applications are typically processed within 1–2 working days” (by HDRUK).

The gateway login page[1] tells the entire internet that it can be accessed in various ways:[2]



The gateway allows anyone with a Microsoft (Azure) account – including a free consumer account – or an account at any organisation using Microsoft (Office) 365 / Entra ID / Active Directory, to perform arbitrary analyses on the data entrusted to you.

HDRUK says they “reserve the right not to provide access if your Safe People or Safe Project status is indeterminate”. CPRD is not determining the safe people or safe projects – CPRD have abdicated that role and given it to HDRUK who have no obligation to meet

[1] Click “Signin” on the top right of any gateway page - there’s no way to link to it.
[2] We observe HDRUK has decided to take away the “Sign in with ORCID” that was there previously, and HDRUK has currently splits access in ways they previously did not – more HDRUK choices not CPRD choices

CPRD's standards, and can change them at will. That is clearly contrary to what the CPRD website says about who makes decisions about CPRD data.

HDRUK and CPRD acknowledge numbers leave the system, and those numbers are the problem.

If a person requests access for one project, there is no limitation or management against use for others other than the promise made to HDRUK, a promise HDRUK are in no position to assess or ensure is kept and CPRD are in no position to sanction. Gateway users need not sign the CPRD agreement, and have no contractual obligation around ongoing payment for access. Additionally, if HDRUK rejects or terminates one account, the user can simply create another. Would CPRD even be informed?

The UK Data Service defines "Safe projects: research projects are approved by data owners for the public good." – HDRUK is not the owner of CPRD data.

Indeed, when HDRUK says "Safe People: Only approved users that have a verified institutional email address and appropriate roles within those organisations can access Cohort Discovery", that is not the definition of Five Safes used elsewhere. The only technically "verified" criterion of HDRUK is that the user has an institutional email address (which isn't necessarily the email associated with the login).

Professor Felix Ritchie's "FiveSafes.org" defines "Safe people – can the user/s be trusted to use the data in an appropriate manner?", yet CPRD and HDRUK have no idea of the answer to that question.

HDRUK does not and cannot uphold CPRD's ban on "rate estimation"

Discussing your "Exemptions and analyses for medicines regulatory bodies and/or public health agencies", the CPRD website says:

> "Analyses permitted under this exemption are limited to numerator or denominator counts only as outlined below. Rate estimation and other analysis including hypothesis testing are not permitted"

Such a restriction does not appear in HDRUK's process, and neither HDRUK nor CPRD is able to enforce one.

Numbers leave the system, and the system has no idea what those numbers mean.

The HDRUK gateway has a fundamental, reckless and known design flaw

medConfidential has long argued that the "cohort discovery" feature is dangerous and reckless by design – we see no way it can be made safe.

The feature, built under the project name "co-connect"[3] to an HDRUK requirement, was designed to allow easy querying of data. Integration with the gateway has not solved any of the fundamental problems with the approach, and HDRUK has not changed the design.

3

Requiring a member of staff to swipe right on every applicant is not a reliable mitigation. Someone glancing at an application was insufficient for UK Biobank and cannot be sufficient for CPRD. You don't even ask the one question UK Biobank ensured it always received a confirmed and validated answer to: "*have they paid*?". The supposed mitigations simply repeat UK Biobank's approach to access 'control', which is to allow as much use as possible.

It is also unclear whether, if the gateway were to revoke permissions, queries would continue for that login session – because HDRUK tells users "If your access has recently been approved, you might need to sign out and sign back in to your Gateway account to refresh your permissions before continuing." This does not inspire confidence.

The gateway now lets anyone with access count patients with (not quite) arbitrary criteria:

- "the number of female asthmatics under the age of 35"
- the number of female asthmatics under the age of 36?
- the number of female asthmatics under the age of 34?
- the number of male asthmatics under the age of 35?
- the number of female white asthmatics under the age of 35?
- the number of female black asthmatics under the age of 35?
- the number of female black asthmatics with HIV under the age of 35?
- the number of pregnant female black asthmatics with HIV under the age of 35? etc

Subtracting one figure from another isolates ever narrower groups, eventually individuals. Is there any review of the various queries that can be put together, and published within permissions HDRUK give their users over CPRD data? You will be aware that any review can only take place when the attackers already have all the answers they desire.

We've written on the co-connect project at various stages of its work.[4] Reflecting HDRUK's shared culture with UK Biobank, HDRUK had decided what they wanted to do irrespective of what they learnt after making the decision, and ignored all feedback that might have nudged them towards a better approach.

We should note the contracted research team who built Co-Connect did engage at the time – they understood that the problem was not the software they had built, but the specification they were contracted to deliver against — which they could not change, and which is where all the problems continue to lie..

All our 2022 criticisms basically still stand

HDRUK (privately) recognised our criticisms of their approach, but they pushed ahead because that is what HDRUK leadership had already decided to do. The parallels to UK Biobank continue.

HDRUK shares a culture with UK Biobank, and now seems to share a project review philosophy with them. What's CPRD's excuse?

---

[4] Software and Context from 2022. Unanswered questions incorporated into wider questions about HDRUK. We would also draw your attention to the Office of Statistical Regulation response to a question on the project: https://www.youtube.com/live/TETh7aUfky4?si=f-vAQezsvSlKQLHm&t=2653

Yours sincerely,





Phil Booth, medConfidential Sam Smith, medConfidential

cc: HDRS, DH/E, JGPITC, Prof Felix Ritchie

PS – In writing this letter, we saw that the CPRD approval process has categories for:

> "3. Ministerial Requests i.e., research in support of security, intelligence, prosecution, international relations
> 4. Research requested by law or court order"

How many requests have come in each of those (sub-)categories, and are they reviewed by your oversight processes? Where are they publicly disclosed? Is this "Ministerial Requests" (which can be political) or "intelligence, prosecution" use what GPs gave CPRD for?