Developing a methodology to assess the consistency of fitness to practise outcomes

Prepared for the Professional Standards Authority
By the Research Department of Medical Education
UCL Medical School

UCL Medical School
RDME
Research Department of Medical Education

Report prepared by:
Professor Ann Griffin
Dr Asta Medisauskaite
Dr Shah-Jalal Sarker
Dr Rowena Viney
Dr Laura Knight
Dr Judith Tweedie

Contents

Abbreviations

Fitness to Practise (FtP)

General Chiropractic Council (GCC)

General Dental Council (GDC)

General Medical Council (GMC)

General Optical Council (GOC)

General Osteopathic Council (GOsC)

General Pharmaceutical Council (GPhC)

Health and Care Professions Council (HCPC)

Health and Care Professions Tribunal Service (HCPTS)

Nursing and Midwifery Council (NMC)

Pharmaceutical Society of Northern Ireland (PSNI)

Medical Practitioners Tribunal Service (MPTS)

Professional Standards Authority (PSA)

Executive summary

Background

The drive for consistency across healthcare regulators is critical to ensuring fairness and justice for professionals, for patients, and for maintaining public trust and confidence. Concerns have been raised about inconsistency in the fitness to practise (FtP) proceedings of the health and social care regulators. In 2018 a rapid policy review into gross negligence manslaughter in healthcare chaired by Professor Sir Norman Williams recommended an investigation of "*the extent and reasons for different fitness to practise outcomes in similar cases and, if appropriate, recommend changes to ensure greater consistency*".(1) The Professional Standards Authority (PSA) has been commissioned by the Department of Health and Social Care to address these recommendations together with an academic partner. As the first step towards a full investigation, this scoping review has examined the ways by which consistency could be researched and what sources of data are available to examine these issues. This report of the scoping review presents its findings and concludes with a proposed comparative methodological approach (evaluation framework) to assess consistency in FtP procedures and outcomes across the nine health and social care regulators in the UK.

Research questions

As set out in the original tender this scoping review aimed to answer the following research questions:

Research Question 1: What are gaps for research in the existing literature?

Research Question 2: Are the potential sources of available data sufficient and veracious to analyse the consistency of fitness to practise outcomes?

Research Question 3: What methodological approach would help to assess consistency in the context of a full study? What are the strengths and limitations of all parts of such a methodological approach?

Research Question 4: What is the scope of a potential full research project, including costings, timelines, etc.?

Methods

For this scoping review we performed a literature review and investigated the potential sources of data which could be used to examine consistency. This included: informal interviews with seven health and social care regulators, an analysis of publicly available data on FtP including regulators' annual reports and investigating the PSA's section 29 database. We drew on all this information to explore what is known about consistency and what data is available for future analysis, in order to inform the methodological approach of a future research project which is needed to answer the following questions:

Full study- Research Question 1: What are the broad points of consistency and difference within the current fitness to practise procedures of the nine regulators?

Full study- Research Question 2: What is the potential impact of different variable factors within fitness to practise procedures?

Full study- Research Question 3: What is the potential for different academic disciplines to contribute to addressing the question of consistency, in light of the range of variable factors?

Full study- Research Question 4: How might legal representation of registrants affect fitness to practise outcomes?

Results

We found limited research investigating consistency in FtP in health and social care and, of import, there were no academic studies analysing consistency across the different UK health and social care regulators. Some studies analysed within-regulator consistency and most of the studies analysed FtP processes in medicine.

Informal interviews with regulators revealed a common four stage process to investigating registrants with FtP concerns, however, there was a significant variation in underlying practices and processes including those related to: allegation categories, outcome categories, nomenclature, legal frameworks, decision-making staff, and nature and type of panel (including location of hearing). Regulators typically had electronic case management systems to store data but the datasets contained a limited range of registrant's sociodemographic details. Registration databases were reported to contain more detailed data. Qualitative data was mostly available across all four investigative stages; this dataset could be sizable and contained sensitive data, however data at the earliest stages of investigation was naturally more limited. All regulators interviewed stressed that any future research would need to have a clear rationale. Data sharing and resource issues would need to be individually negotiated with the regulators. The PSA section 29 database contains all the FtP cases that reached the hearing stage, with an estimated size of 22,000 cases. It is a central site for determination documents. However, it would be insufficient for statistical examination because of coding problems and missing data.

In conclusion, the data currently available for examination of consistency is limited. Of note, registrant's sociodemographic details are not currently fully recorded and exactly what is recorded differs between regulators. Significant time will be required in preparing what data is available for statistical analysis and many important variables will be missing.

Case files contain a wealth of information, especially towards the end of an investigative process when they can reach 1,000s of pages. However, they contain sensitive information and so will require redaction as well as requiring significant time to analyse.

Key methodological factors to consider for future research

- The myriad factors that influence consistency at **the level of society, the organisation, and the individual** (see Figure 2, p. 27);
- There was **variability between regulators** in terms of FtP processes including differences in investigative process at the four stages, legal frameworks, categorisation of allegations and outcomes, databases not readily available for statistical analysis, and less detail for cases at earlier stages of the investigation;
- **Individual variability** in terms of the various professional standards, sociodemographic details, number of registrants, allegations categories and confounders, for example multiple allegations;
- The need for comparative analysis across nine regulators with **variable number of registrants** in the FtP process.

Methodological approach

Despite these obstacles, a robust and comprehensive programme of research is possible. Research needs to examine the regulatory contexts, the procedural elements, individual registrant factors and consistency as it relates to both, outcome measures and decision-making. Future research in to

consistency needs to not only see if consistency exists, or not, but should provide evidence for the causal factors for it; in this way research can recommend changes so that any inconsistencies can be addressed. Research of this nature needs to be highly rigorous because of its implications for regulators and registrants, and most importantly for patient safety.

Synthesising the findings of this scoping review has informed the development of an evaluation framework based on all the factors known to impact consistency. The scoping review suggests that future research examine three structural levels which influence consistency: the wider influences at the Macro level, the organisational issues at the Meso level, and individual factors at the Micro level.

A mixed method programme of research will be vital for understanding consistency in health and social care regulators' FtP proceedings. Crucially any research undertaken must explore any associations with individual factors, contextual factors, and generate causal explanations.

Structure of the report

The following sections of this report are structured according to the scoping research questions and in each section after discussing the main findings we highlight their implications for future research. We begin with Scoping question 1: What are gaps for research in the existing literature?

Scoping question 1: What are gaps for research in the existing literature?

The aim of the academic literature search in this report was to identify what was already known from the academic literature about consistency between health and social care regulators and what critical knowledge gaps could be filled by future research.

Methods

We have conducted a scoping literature review of existing academic literature, searching for relevant studies in the ProQuest database. We have specifically selected this database as it includes a large number of journals that focus on the study of healthcare professions. We searched for:

1. Empirical studies (qualitative studies, quantitative studies, systematic reviews) and excluded such review papers as commentaries and editorials;
2. Studies on health care professionals and not undergraduate students;
3. Studies conducted about UK health care professionals;
4. Studies about various aspects of FtP procedures, e.g. allegations, procedures, outcomes.

Results

We used "fitness to practise" as the key word to search for relevant papers. It revealed 1699 hits. After reviewing the titles and abstracts (where applicable) of these 1699 hits, we identified 19 relevant studies. These studies were published from 2009 to 2019. Twelve studies were about doctors; four about dentists; one about nurses; one about pharmacists; and one about social workers. Studies used various research methods and had various aims:

a) **Systematic review** – one study about doctors from various countries including the UK.
   Aim: to examine the association between gender and poor performance in doctors;

b) **Policy documents analysis** – one study about nurses.
   Aim: to examine the content of Higher Education Institutions' FtP policies and to compare this with the regulator's policies;

c) **Survey** (quantitative and qualitative analysis) – three studies, all about doctors.
   Aims: to investigate the impact of complaints on doctors' psychological well-being and health;

d) **Case analysis** – seven studies about doctors, dentists, pharmacists, and social workers.
   Aims: to describe cases (the prevalence, allegation types, outcomes, etc.); to analyse the process (role of reflection in continuous education; analysis of consideration of aggravating and mitigating factors; decision-making); to make predictions (if risk factors at medical school relate to subsequent professional misconduct);

e) **Database analysis** – seven studies about doctors.
   Aims: to describe complaints (allegation types, outcomes, etc.); to analyse registrants' characteristics (primary qualification, gender and outcome); to predict outcomes (from exam results, e.g. professional and linguistic assessments board (PLAB), postgraduate).

Please find a summary of the 19 relevant studies in Table 1 below.

Table 1. Summary of studies on fitness to practise in healthcare

Author	Year	Research methods	Prof group	Data	Key findings
Singh et al. (2)	2009	Case analysis	Dentists	209 cases (publicly available; 5 year period) brought before the Professional Conduct Committee of the GDC	• An annual increase from 23 cases in 2003 to 65 cases in 2007 (0.18% of registered dentists); • Over half of the registrants (123, 58.9%) were UK graduates; • Most charges are related to clinical practice: poor treatment (83), poor records (61), radiation (59) and NHS fraud (57); • Outcomes: 56 erasures and 37 suspensions.
Yates & James (3)	2010	Case analysis	Doctors	59 UK medical graduates (1959-1997) with professional misconduct in GMC (1999-2004) and 236 controls	• Male sex, lower social class, and poor early performance (failure of early or preclinical examination) at medical school were independent risk factors for subsequent professional misconduct.
Humphrey et al. (4)	2011	Database analysis	Doctors	7,526 inquiries to the GMC concerning 6,954 doctors	• 30% of inquiries at the initial triage concerning UK graduates had a high impact decision vs 46% concerning international medical graduates; • 5% of inquiries at the investigation stage concerning UK graduates were referred for adjudication vs 10% concerning international medical graduates; • 1% of inquiries at the adjudication stage concerning UK graduates led to erasure/suspension vs 4% concerning international medical graduates.
Unsworth (5)	2011	Policy analysis	Nurses	44 Higher Education Institution policies compared with HCPC policies	• Many Higher Education Institutions' policies have significant gaps; • Some Vice Chancellors have the power to overturn FtP decisions on appeals.
Campbell et al. (6)	2013	Database analysis	Doctors	Complaints about 105 anaesthetists in 2009 (GMC dataset)	• Compared with doctors in general, anaesthetists had a lower rate of referral; • Compared with doctors in general, anaesthetists were less often referred by members of the public.
Unwin et al. (7)	2014	Database analysis	Doctors	329,542 doctors on the register in 2013	• Female doctors had nearly a third of the odds of having sanctions compared to male doctors; • Female doctors who had specialised as general practitioners were the least likely to have sanctions against their registration.
Donaldson et al. (8)	2014	Database analysis	Doctors	The National Clinical Assessment Service referral data (6,179 doctors; 2001–2012)	• The annual referral rate was five per 1,000 doctors; • Non-UK graduates were more than twice as likely to be referred as UK graduates; • Male doctors were more than twice as likely to be referred as women doctors; • Doctors in the late stages of their career were nearly six times as likely to be referred as early career doctors.

Bourne et al. (9)	2015	Survey	Doctors	Anonymous survey: 7,926 doctors with recent/current, past, or no complaints	• Doctors with current/recent complaints were two times more likely to report suicidal ideation or self-harm thoughts; • After whistleblowing 20% felt victimised, 38% felt bullied, 27% spent over one month off work.
Unwin et al. (10)	2015	Systematic review	Doctors	Literature review	• 27 out of 32 reports found that male doctors were more likely (from 19 studies nearly 2.5 times the odds) to have experienced medico-legal action than female doctors.
Gallagher et al. (11)	2015	Case analysis	Pharmacists	51 cases (requested from the GPhC)	• There is a link between both risk of harm and dishonesty as aggravating factors and erasure; • The GPhC considers relevant aggravating and mitigating factors at all stages of their deliberations into practitioner misconduct.
Gallagher & De Souza (12)	2015	Case analysis	Dentist	66 committee transcripts (publicly available)	• The aggravating factors correlated with erasure; • The GPhC considers relevant factors at all stages of their deliberations into practitioner misconduct.
Bourne et al. (13)	2016	Survey: qualitative study	Doctors	Anonymous survey: picked 100 from 3,417 doctors answers	• Many doctors felt that the complaint was unfair, felt unsupported and fearful of the consequences; • The most stressful parts were the prolonged duration and unpredictability of procedures, poor communication, managers' incompetence, and perceiving that processes are biased in favour of complainants; • Many were considering changing career after a complaint.
Brindley (14)	2016	Case analysis	Dentist	56 cases (publicly available) which resulted in conditions	• For 50 cases, registrants were asked to produce reflective logs; • 38 registrants were asked to submit personal development plans for regular reviews.
Bourne et al. (15)	2017	Survey	Doctors	Anonymous survey: 7,926 doctors with recent/current, past, or no complaints	• While 61% of doctors felt supported by colleagues, only 31% felt supported by management; • Perceived support by management, speaking to colleagues, fair/accurate documentation, and being informed about rights positively links to well-being and reduced defensive practice; • Following complaint doctors worried most about professional humiliation (80%).
Tiffin et al. (16)	2017	Database analysis	Doctors	Data linkage study: 27,330 international medical graduates	• Lower PLAB part 1/2 scores at first sitting, multiple attempts at both parts of the PLAB, lower international English language testing system (IELTS) reading and listening scores and higher IELTS speaking scores predicted censure (versus no censures or referrals). • The limit of four attempts at both parts of the PLAB would reduce the risk by only approximately two censures per 5 years.

Mehdizadeh et al. (17)	2017	Database analysis	Doctors	1,111 doctors who had a performance assessment between 1996 and 2013	• The rate of performance assessment varies by place of medical qualification and by year; • Non-UK graduates were more likely to have a performance assessment than UK graduate doctors, except for South African trained doctors.
Neville (18)	2017	Case analysis	Dentists	6 cases (publicly available) related to social media (2013-2016)	• 2.4% of all GDC cases were related to breaches of the social media guidelines; • Most cases were against dental nurses; • Most common type of complaint was inappropriate Facebook comments.
Leigh et al. (19)	2017	Case analysis	Social workers	34 cases (publicly available): hearings about performance issues	• The seriousness of the registrant's misconduct or competence does not necessarily relate to the severity of sanction; • Significant factors in decisions on sanctions: registrant's engagement with the process, registrant's insight into the issues and credibility as a witness.
Wakeford et al. (20)	2018	Database analysis	Doctors	Data linkage study: 27,562 MRCGP AKT; 17,127 MRCGP CSA; 36,934 MRCP (part 1); 28,011 MRCP (part 2); 26,752 MRCP (Paces)	• Knowledge and clinical assessments predict future FtP sanctions; • Doctors on the 2.5th percentile of postgraduate exam performance were about 12 times more likely to have FtP problems than those on the 97.5th percentile.

Conclusion

The academic literature scoping review revealed a limited number of studies on FtP. Regarding consistency, there were no studies analysing consistency across different UK health care regulators. Some studies analysed consistency within a regulator, for example whether the regulator considers aggravating and mitigating factors at all stages of their deliberations into misconduct. The findings, however, are limited as such studies are often undertaken as a quality assurance investigation or audit.

The vast majority of studies (12 of the 19 included in this review) were about doctors and we found a limited number of studies on dentists, nurses, pharmacists, and social workers. We did not find studies on chiropractors, osteopaths, or opticians as well as other professionals regulated by HCPC. In addition, many studies were either descriptive in nature (e.g. discuss a spread of outcomes) or analysed what factors link to FtP outcomes (e.g. whether postgraduate exam results link to a doctor being sanctioned).

The lack of studies on consistency of FtP processes and outcomes across the healthcare regulators confirms that research in this area will be of significance.

Implications for future research

Based on our review, we conclude:

- There is a **lack of high quality** studies about FtP;
- Of the limited studies available **many focus on FtP outcomes**, with just a few studies on processes;
- There are limited studies on **professionals** other than doctors.

Scoping question 2: Are the potential sources of available data sufficient and veracious to analyse the consistency of fitness to practise outcomes?

Information was synthesised from three sources to produce this section of the report: 1) informal interviews with seven regulators about their FtP procedures and available data, 2) publicly available data on regulators' websites including annual reports, and 3) analysis of the PSA database. The first section of the report presents results from informal interviews with the regulators. The second section discusses the PSA section 29 database and the last section specifically deals with data currently available to understand the impact of legal representation on consistency. Each of these three sections concludes by discussing these findings and their implication for future research.

1. Interviews with the regulators

The aim of informal interviews with health and social care regulators was to explore the sufficiency and veracity (data quality) of their available sources of information (databases and documents) for assessing consistency in the context of a full study. On 14-16th February all nine UK health and social care regulators were invited to take part in informal telephone interviews. A checklist was developed to guide and standardise the interviews (Appendix A). This checklist was sent in advance of the phone interviews to ensure the fullest data capture. Two "follow up" reminder emails were sent to non-responders when required.

The data-gathering phase ended on the 31st March 2019, at which point seven of the nine regulators invited to take part had engaged. Of these seven regulators, five regulators took part in informal telephone interviews; one regulator partly completed the interview but subsequently sent their responses to the rest of the checklist questions via email; and one regulator was not able to participate in an informal interview but sent the completed checklist via email. Some interviewees' roles within the organisation meant that they were not able to answer some of the questions in detail.

The following sections detail the results with regards to:

1. Allegation categories and FtP outcomes;
2. FtP processes;
3. Quantitative databases;
4. Qualitative data available;
5. Other relevant documentation;
6. FtP hearings;
7. Data retention policies;
8. Data sharing in future research;
9. Implications for future research.

Allegation categories and FtP outcomes

The number of high-level allegation categories vary across the regulators. One regulator had seven allegation categories, one regulator had six, three regulators had five, and two regulators had four. However, in six of the seven regulators the high-level allegation categories were broadly similar and could potentially be grouped as:

- Misconduct;
- Lack of competence;
- Conviction or caution for criminal offence;
- Health;

and

- Determinations by other health or social care organisations;
- Not having the necessary knowledge of English.

The categorisation and range of the investigative outcomes at both the case examiner/investigative committee (stage III, see below) and hearing stage (stage IV), also varied across the regulators. For example, case examiners/investigative committees had between two and eight outcome options. This range also highlights the procedural variation according to the regulator.

Similarly, there was some variation in the categorisation of the FtP hearing outcomes but this was fewer than case examiner/investigative committee outcomes, with regulators having between four and six outcome options.

Broadly, hearing outcomes map to:
- No further action;
- Advice;
- Caution, warning or reprimand;
- Conditions;
- Suspension;
- Erasure.

FtP processes

All regulators follow the same four-stage process for FtP, which is:
1) Complaint (stage I)
2) Triage (stage II)
3) Investigation (stage III)
4) Hearing (stage IV)

There were however differences in the processes underpinning the four FtP stages across the regulators. Nomenclature for the stages in the FtP process varied across regulators. Stage II also varied in process and is undertaken by personnel with different roles within the organisation. Of particular note, at stage III of the process four regulators used case examiners, four used an investigating committee, and one used a scrutiny committee with variations in ability to impose sanctions.

All regulators who responded to the research team confirmed that information is captured at all stages of the FtP process. The data collected and the method used to store data varied across the regulators; however, information stored in case files was, on the whole, consistent. Six out of the seven regulators held information gathered at stage I, typically detailing the nature of the complaint and initial information. Six out of the seven regulators gathered data for stage II (triage/screening) and this typically included registrants' personal details and details of the complaint, including allegation category. All regulators held data for stage III (case examiner/investigation), most of which was the same data gathered from the earliest stages of the complaint but updated as new information arrived. At stage IV (hearing) more detailed data was available, including details of the hearing itself.

Quantitative databases

Complaint details

All regulators who participated in this project retain data regarding the FtP processes. The majority of regulators reported that older cases consist of less rigorous and comprehensive data, however they reported an improvement in the quality and extent of data in more recent years. Four out of seven regulators reported significant changes to FtP databases since data collection began. No regulator has

a statistical database but all regulators report using either a case management system or excel spreadsheets. The number of cases held by a regulator varies as a reflection of the number of registrants, ranging between 9000/year and 33/year. There was variation in how many years data was available for, ranging from 4 to 17 years. All regulators participating in the project reported collecting data for 1) duration of case; 2) enquiry source; 3) allegation category, and 4) outcome.

Allegations can, and frequently are, added to a case as the investigation proceeds. Furthermore, within the databases it is not possible to decipher which allegation had the greatest influence or led to the outcome decision. One regulator reported that their FtP database held information on decision-making, whereas all other regulators reported outcomes were available but not information regarding the decision-making processes.

Databases differ in how a complaint/FtP is recorded. For the majority of regulators each data line in the database represents a complaint/concern/allegation against which the case is held. Databases can be interrogated for multiple complaints associated with a single registrant.

Demographic information

Regulators reported that registrants' details were held on separate registrant databases. For FtP databases, information recorded tended to vary across the regulators (examples presented in Table 2).

Table 2. Percentage of sociodemographic characteristics held by regulators on FtP databases

Gender	Place of qualification	Ethnicity	Age	Time since qualification	Organisation	Job title	Status
85%	42.8%	28%	71%	14%	14%	0%	0%

Qualitative data available

All regulators participating in this project hold case files for FtP processes. One regulator reported a mixture of paper and electronic files, all other regulators store case files electronically with data scanned and uploaded as required. Most regulators keep relevant documentation for each stage of the process. Typically, this was described as including all material associated with the case such as phone calls, emails, letters, witness statement, statement of case, internal communication, health records, assessments and determinations.

There was no average length for case files and some were several thousand pages long. Cases closed at complaint stage tend to be shorter in length although this can be variable. The format for case files also varies across regulators. One regulator reported holding two types of case files which can be active at the same time 1) triage, investigation and case examiner stage; and 2) legal case files.

All case files were reported to contain identifiable information pertaining to the registrant, patient and frequently other members of the healthcare team. One regulator had files redacted as part of a research project, all other regulators stated that redaction would be necessary for future research.

Other relevant documentation

Other documentation that interviewees suggested may be useful for future research include:

- Internal reports on quality assurance of decision-making;
- Internal guidance documentation to improve consistency in investigations;
- Post-hearing feedback;
- Internal guidance for decision-making;
- Operational guides that ensure public decision-making guidance is fully embedded in processes;
- Training videos;
- Staff manuals;
- Audits of work and processes;
- External research by researchers;
- Regulation manual.

FtP hearings

Most regulators audio recorded their FtP hearings. Transcripts can be requested in specific situations but are not routinely available. Determinations are published publicly but may be partially redacted containing the registrant's name. Private determinations may contain more information than public, including more sensitive data. Determinations contain information about the decision-making process. Most, but not all, regulators include determinations within case files if the hearing has concluded.

The hearing panels typically consist of one lay, one professional, and one other member. Guidance for panel membership pertains to professional/lay representation as opposed to ethnic or gender diversity. Legal assessors are usually present at the hearing and may sit on the panel in some instances. The use of a legal chair is typically limited to one regulator. Four regulators reported using guidance for the recruitment to panel.

Three regulators have separate professional conduct and health committees and one has three committees consisting of a conduct committee, health committee, and professional practice committee. Three regulators use geographic locations across the UK for hearings.

Data retention policy

The retention policy of each regulator governs how long data is held for and therefore what data is available for future research. The retention policy is dependent on what type of data is held, what has been published, the nature of the case, and the outcome. For example, in one regulator time limits are in place for how long case files are held for:

- Case files held for two years if case closed at triage;
- Case files held for five to ten years if closed at investigation;
- Case files held for over ten years if case proceeded to tribunal.

At a different regulator case information is held for twenty years for cases closured at triage.

Data sharing and future research

All regulators were clear that future research would require a data sharing agreement with each individual regulator and be fully compliant with General Data Protection Regulation (GDPR). The rationale of the research and its scope would be clear. The research proposal should include information as to any potential impact on registrants. It is likely to require agreement at executive or

council level in each organisation. Data shared would be required to be as specific and proportionate to the needs of the project as possible.

All regulators participating in this project described redaction as a significant resource in terms of cost and personnel. The majority of regulators stated they would not be able to complete this 'in-house' with current resources. It is possible to receive transcripts of FtP hearings but this would incur considerable transcription costs.

Implications for future research

- **Allegation categories** – there is significant variation in categories used by regulators based on our interview data and our analysis of publicly available regulator FtP reports. The variation in categorisation should be a consideration on any future research as it affects regulators' reporting, internal policies, and guidance for registrants.(21) A common typology will need to be devised to group categories for research purposes.
- **Multiple allegations** are often accumulated over the course of the investigation and will have an impact on the outcomes. In order to account for this we suggest that future research should take into account multiple allegations in their analysis. For a more detailed discussion about such analysis, please see p. 37-38.
- **Outcome categories** – there is less variation in the outcome categories across the regulators; however, variation still exists and will need to be reviewed in order to have a comparable set of outcome categories applicable to all regulators. Further quantitative data will be needed to explore this in more detail.
- Whilst there is a commonality in the four-stage **process of FtP investigation,** there is variation between the regulators about nomenclature, decision-making personnel, processes, and legal frameworks. It is therefore important to have a detailed insight into this variation in order for conclusions to be drawn about its meaningful impact on consistency or not. This will need to be explored both qualitatively and quantitatively.
- All regulators retain FtP data in electronic form. **Quantitative databases** are case management systems and not statistical databases. Only selective variables from case files and other sources like determination documents/case proceedings are entered into the electronic case management system, which is not consistent across the regulators. Databases have gone through significant changes since data collection began and hence more recent databases (past few years) are more reliable. Future research using regulators' databases needs to be mindful of these facts and allow enough time for combining and cleaning datasets.
- Regulators normally **collect data** on 1) duration of case, 2) enquiry source, 3) allegation category, and 4) outcome; information regarding the decision-making process are not usually in the regulators' FtP database. Demographic data, which may be significant in a future exploration of consistency, is variably included on the current datasets. Demographic data from case files or registrants' databases could be used to supplement existing regulators databases, crucial for a robust quantitative study. Regulators range from very large to much smaller organisations with proportionally less complaints. Statistically analysing some variables where the numbers are smaller introduces limitations, however these could be overcome by combining statistical with qualitative analysis of case files.
- The differences in data **retention policies** need to be balanced with gaining sufficient quantitative data in order to robustly examine the variables likely to be of significance in a

study of consistency. It is likely that any future study would need to examine several years' worth of data and preferably more recent data, as, according to these informal interviews, it is likely to be of a higher quality and more complete.

- **Case files** are an important source of information to explore the consistency of decision-making. They are more likely to contain detailed individual demographics, information about legal representation, and the nature of the complaint including the outcome. However, case files contain identifiable data and only two regulators reported the possibility of anonymisation. Significant barriers to using case files include the cost and time of redacting them.
- **FtP determination documents** are an important source of exploring decision-making with regard to consistency. These are available publicly, but private determinations tend to have more detail. If not redacted it will be associated with the higher costs.
- **Transcripts of hearings** will be a useful source of very detailed information and be able to explore causal factors for various outcomes including the influence of legal representation. As not all hearings are transcribed, and some hearings can take up to five days, time spent on transcribing and analysing hearings would have resource implications.
- **Other documentation** (e.g. internal reports) may provide important insights into procedural variation between the regulators.
- **Regarding ethics/data protection/research governance** we do not anticipate issues with ethical approval for future research. However, regulators will need confidence that data protection and GDPR issues have been addressed. Individual data sharing agreements need to be reached with each regulator, but any future research needs to ensure the highest degree of research governance and data security. It is recommended that this be done on a data safe haven[1].
- Within the **timescale** of the research only seven out of the nine health and social care regulators were able to take part in informal interviews. Future research needs to be mindful of the variation in regulators' resources and capacity to take part in further study. In addition, working closely with the regulators is important as regulators are experts in their FtP investigations (e.g. they can explain any anomalies) and data holders (e.g. provide information which researchers need). Future research needs to first establish a good working relationship with regulators which is essential for the success of this project.

2. The PSA section 29 database

The PSA section 29 (S29) database holds reports on all FtP cases where a decision has been made, recording all cases that have completed the hearing stage of the FtP process (stage IV).

Quantitative data

The S29 database was reported to contain around 22,000 cases from 2013 to date and includes cases from all the nine regulators. At present, cases can be searched based on: registrant name, registrant gender, regulator type, allegation type, outcome, PSA appeal, etc. The research team at UCL audited the database in March/April 2019 to inform decisions about its use for any future research.

[1] A data safe haven is a secure portal for storing research data.

The S29 database can be searched to find similar cases, e.g. in terms of allegation type, outcome, and registrant characteristics. For example, Table 3 shows the distribution of cases according to four key hearing types (*note*: this table does not include results for other hearing types that are available in the search: language impairment, non-compliance, and restoration cases). There is a wide variation across regulators in the number of cases recorded in the four hearing type categories. Furthermore, these types also have many lower level allegation sub-categories (charge summary: e.g. sexual misconduct, treating without consent, verbal abuse, etc.). Therefore, any future sample should be chosen carefully to account for this variation and ensure adequate representation across the regulators.

Table 3. Total cases in the database according to hearing type and regulators

	Regulators								
Hearing type	**GMC/ MPTS**	**HCPC**	**GOC**	**GCC**	**GOsC**	**NMC**	**GDC**	**GPhC**	**PSNI**
Health	456	214	20	-	6	1444	197	194	-
Conduct	1447	2339	194	85	132	8876	1328	589	20
Conviction/ caution	277	282	15	7	6	753	122	144	2
Performance	212	620	3	1	9	1066	129	29	-

Note. Totals: GMC/MPTS=2042; HCPC=3026; GOC=205, GCC=90, GOsC=143, NMC=11400, GDC=1480, GPhC=748, PSNI=20. Totals are lower because some cases are classified under more than one hearing type.

The same situation is apparent for allegation outcomes as Table 4 illustrates. Table 4 shows the distribution of cases in the S29 database according to five key hearing outcomes (*note*: outcomes for four hearing types presented in Table 3). It is not presented in this report, but it is possible to search for other outcomes in the database, e.g. financial penalties, conditions after suspension, restoration, etc.

Table 4. Total cases in the database according to outcome and regulator

	Regulators								
Outcome	**GMC/ MPTS**	**HCPC**	**GOC**	**GCC**	**GOsC**	**NMC**	**GDC**	**GPhC**	**PSNI**
No impairment	424	810	74	50	59	2716	465	192	4
Caution/warning	55	232	23	19	36	691	93	29	-
Condition/undertaking	265	345	22	8	12	2045	236	78	3
Suspension	680	764	38	6	18	2921	471	291	4
Erasure/striking off	445	789	41	6	17	2420	154	139	10

Note. Totals: GMC/MPTS=2042; HCPC=3026; GOC=205, GCC=90, GOsC=143, NMC=11400, GDC=1480, GPhC=748, PSNI=20. Totals are lower because some cases are classified under more than one hearing type.

The database contains a number of missing values. For example, Table 5 shows the missing values for gender.

Table 5. Total cases in the database according to gender and regulators

	Regulators								
Outcome	**GMC/ MPTS**	**HCPC**	**GOC**	**GCC**	**GOsC**	**NMC**	**GDC**	**GPhC**	**PSNI**
Total cases	2213	3092	216	92	146	11900	1539	768	22
Gender									
Male	1532	1484	143	54	93	2907	923	447	10

Female	340	1527	51	13	44	8345	442	219	11
Not specified	16	-	2	-	-	25	1	2	-
Missing	**325**	**81**	**20**	**25**	**9**	**623**	**173**	**100**	**1**

Some variables were added by the PSA to the database more recently and therefore there are a limited number of cases that can be screened based on these newer variables – for example, the legal representation variable was added around 18 months ago. It is not possible to search the database on registrants' characteristics such as age, place of primary medical qualification, time since qualification, and other similar more specific characteristics.

The database appears not to handle searches with large numbers. Despite knowing that it contains 22,000 cases only 20,100 cases were reported in our search. Whilst we are aware that data can be exported into software programmes which would alleviate some problems, the issues with limited coding and missing data limits its use for further quantitative analysis.

The PSA are planning to update the S29 database in October, 2019 to include more features which may increase its utility for future research.

Qualitative data

The variety and quality of the documents included in the S29 database vary. The PSA only requests detailed documentation to examine (and store on the S29 database) if the determination of a case is considered to be too lenient. Documents in these types of file include: evidence bundles, determination documents, hearing summaries, and transcripts. For cases where the determination is not a concern, case files only contain determination documents and email exchanges. It is important to note that because erasure is the most severe sanction, the cases with such an outcome are not investigated by the PSA (as this determination cannot be made *more* severe). This means that detailed information about the decision-making processes of the most severe determinations is not available in this database.

Implications for future research

- **The S29 database holds data only for cases that went to a hearing**. Future research should seek more information from regulators on cases which did not reach this stage of the FtP process.
- In cases where there is less information, when **the PSA is not looking further into a case**, **only the determination documents** could be used for comparisons across regulators.
- The S29 database will have limited use for quantitative research purposes because of a number of **missing values and the limited number of search variables**. Even though it is useful to explore the cases, future research would need to focus on qualitative analysis of cases and/or consider coding cases for quantitative analysis.
- On examining the database, it was found that **the content per case is variable** due to how much detail was requested from the regulator for the PSA to examine. The database contains a range of data that could inform a potential qualitative analysis. However, in order to be able to examine consistency across the nine regulators, future research should consider that the selected data would need to be comparable, and therefore a qualitative analysis would be dependent on how consistent the types of data are across the cases in the S29 database. Cases being scrutinised by the PSA have more data available for **future qualitative analysis:**

- ***Determination documents**:* from the determination documents it would be possible to explore how allegations, outcomes, and decision justifications are described by the various regulators. Discourse analysis (22) would be a useful mode of analysis to examine the language used in the documents, and to identify any issues around power and fairness. The determination documents can also provide some insight into any differences regarding registrants' having legal or other representation. For a more detailed discussion about discourse analysis, please see p. 33-40.
- ***Hearing transcripts***: from the transcripts it would be possible to explore how the different regulators conduct their hearings, including their format, the participants, and whether different aspects of the cases receive different amounts of time within the hearings (see also section on legal representation for more discussion about working with transcripts). Discourse analysis (22) and conversation analysis (23) would be useful approaches to take for this data. For a more detailed discussion about conversation analysis, please see p. 39-40.
- Other types of information such as ***evidence bundles***, while interesting, would be harder to include in a comparative analysis as there is more variability across, for example, regulators and allegation types. Looking at this data could tell us more about the different regulators' consistency in their processes by using them in a more quantitative way, for example by coding the varying types of information provided using content analysis. For a more detailed discussion about content analysis, please see p. 35-36.

3. Data available to understand the impact of legal representation on consistency

Professional regulators' databases

Information about what is currently documented about legal representation was gained through the informal conversations held with seven regulators (see section 1). The practice of actively logging legal representation varies between the regulators. Six regulators actively noted the presence of legal representation, typically on their case management system but sometimes on an external database. One regulator currently does not log this information specifically. When legal representation was logged it was for the final hearing stage investigation. No one listed the type of legal representation used.

However, regulators stored relevant correspondence/documentation on their case management systems as it was presented to them and therefore analysis of these databases could be used to identify the presence and nature of legal representation and at what stage in the fitness practice process it occurred.

The PSA database

The PSA database logs the presence, absence or uncertainty regarding registrant representation on their S29 database. This documentation process started approximately 18 months ago. The PSA database only records stage IV of the FtP process, i.e. hearing stage. The current database differentiates between representation by a lawyer or other representation.

We audited the database on March/April 2019 to inform decisions about its use for any future research. This audit looked at both the quantitative and qualitative data.

Quantitative data

Table 6: Legal representation on the PSA database

Representation	Registrant is represented	Registrant is represented by a lawyer
Yes	1216	1024
No	2140	2318
Unknown	388	401

Table 7. Breakdown of any representation by regulator

Regulator[2]	Any representation "yes" (n=1216)[3]	Any representation "no" (n=2318)[3]	Any representation "don't know" (n=401)[3]
GDC	104	128	34
NMC	553	1315	170
GPhC	47		23
GMC	21	27	13
HCPC	209	538	68
MPTS			67
GOsC			7
GOC			7

Table 8. Breakdown of represented by lawyer by regulator

Regulator	Representation by lawyer "yes" (n= 1024)[3]	Representation by lawyer "no" (n = 2318)[3]	Representation by lawyer "don't know" (n = 401)[3]
GDC	86	144	36
NMC	500	1359	176
GPhC	37		23
GMC	21	27	13
HCPC	105	459	72
MPTS			67
GOsC			7
GOC			7

Table 9. Representation by hearing type category

Hearing type	Any representation (n=1216)	No representation (n=2140)
Health	166	398
Conduct	880	1455
Conviction/caution	143	229
Performance	138	306

Qualitative data

An audit of 30 case files was undertaken to see if they were sufficient for future research examining the impact of representation, including representation by a lawyer, on consistency. Fifteen cases from the database were examined from those who had any form of representation and 15 cases examined from those who were represented by a lawyer. All 30 cases were screened to ensure they were different cases because the filter "any representation" includes those cases represented by a lawyer. The cases spanned the various regulators and a range of allegation and outcome categories. Twenty-

[2] No results were returned for GCC and PSNI.

[3] These numbers were produced by applying the relevant filters on the database. There is a discrepancy between overall totals (yes, no and don't know) and the subtotals generated by individual searches by the various regulators. The reason for this is that there is missing data as not all cases are coded and often the determination documents do not clearly detail this information.

nine of the 30 cases analysed had determination documents available for scrutiny and one case had a full bundle with transcription of the FtP hearing.

*Any representation analysis***:** in the 15 cases analysed evidence for representation was present in only four of the associated documents.

Representation by lawyer analysis: in the 15 cases analysed again only four cases had clearly documented legal representation. In two cases whether representation was present or not was unclear. In one case representation was by a union official and in the other by a mental health worker. In the other eight cases there was no apparent legal representation detailed within the determination documents.

The determination documents generally have sparse reference to the legal representation and any role it plays in the FtP process therefore they would be unsuitable for any further assessment about the impact of legal representation on consistency.

Implications for future research

- Generating statistical data demonstrating any impact of legal representation, particularly at the **earliest stages of the complaint**, will be challenging. Data exists but is incomplete and held in various formats which would need to be collated for future research.
- Using the PSA S29 database will also be problematic because of issues with missing data and coding of the determination documents. **Quantitative analysis** is limited without extensive recoding. There is very little text in determination documents relating to the input of any representation making it inadequate for further **qualitative analysis**.
- Full **case files,** particularly those containing transcripts, would be suitable for further research. The PSA S29 database does contain 67 cases which have been referred to court and contain full case files and could be used for analysis. However, when filtered by "representation present" only 10 cases appeared. This number is unlikely to generate sufficient data to draw robust conclusions.

Scoping question 3: What methodological approach would help to assess consistency in the context of a full study? What are the strengths and limitations of all parts of such a methodological approach?

Introduction

This section of the report begins by reflecting on what is meant by the term consistency. How consistency is understood is important to consider as it will strongly influence the approach to any future research. We then describe concerns about consistency in FtP processes and outcomes amongst health and social care regulators, and discuss the broad range of factors that influence consistency. In order to account for all of the possible causal factors affecting consistency in our suggested methodological approach, we utilise reports and literature in health and social care but also rely on a limited literature from the academic discipline of law.

We then draw on all the findings presented in this report to present an approach by which consistency could be assessed. We take in to account influencing factors that operate at the level of wider societal influence, those that pertain to organisational culture and process, and those individual registrant factors that could impact on consistency – at the level of Macro, Meso and Micro – and use this concept to structure a methodological framework to guide the research.

Addressed in detail are three of the proposed research questions that future research is required for:[4]

> *Full study- Research Question 1:* What are the broad points of consistency and difference within the current fitness to practise procedures of the nine regulators?
>
> *Full study- Research Question 2*: What is the potential impact of different variable factors within fitness to practise procedures?
>
> *Full study- Research Question 4*: How might legal representation of registrants affect fitness to practise outcomes?

What is consistency?

The UK Sentencing Council defines consistency in the following way:(24)

> "*To apply the same purposes and principals of sentencing, and to consider the same types of factors when sentencing… [it's] not "mathematical precision" but the consistency in the application of relevant legal principles*".

The notion of consistency implies fairness or justice; that is similar offences result in similar penalties. According to Krasnostein and Freiberg (25) the drive for consistency places increasing significance on the role of policy and the standardisation of both processes and outcomes.

However, the concept of consistency is not without controversy. Consistent outcomes do not necessarily imply fairness, identical outcomes that have ignored relevant individual and legal factors are equally unjust. Legomsky (26) identifies four broad areas which can result in inconsistency: the incorrect application of legislation, inaccurate interpretation of factual material, inappropriate use of discretionary powers and a mixture of all three areas. Furthermore, similar offences resulting in the same forfeits precludes a more individualistic, case-by-case, assessment of concerns. The drive for consistency can also be viewed as undesirable because it undermines judicial discretion, i.e. the ability

[4] *Research question 3: What is the potential for different academic disciplines to contribute to addressing the question of consistency, in light of the relevant variables?* is addressed in the final section of this report.

of legal persons or panels to take in to account the fullest range of circumstances (individual, contextual, etc.) in making their decisions. Krasnostein and Freiberg (25) suggests that there is "cultural ambivalence" towards consistency in the legal profession:

> "*The tension between individualised justice and consistency is reflected in the potential difference between a sentence based on the circumstances of an individual case and one based on the comparison of similar cases*".

Disparity, the opposite of consistency, is therefore regarded as a necessary feature of an individualistic approach to justice. However, unjustified disparity, either penalties that are too lenient or conversely too harsh, indicates unfairness. The concept of "unjustifiable disparity" should arguably be the guiding principle for future research into this area.

Concerns about inconsistency

In addition to the William's report (1) which we mentioned in our introduction, there are other reviews that observe or discuss possible inconsistencies:

- Inconsistencies in FtP models are discussed in Right Touch Reform.(27)
- Inconsistencies in the proportion of enquiries and complaints (as well as thresholds for investigations) are presented in the GMC UK Health Regulator Comparative Data Report.(28)
- Inconsistencies in the categorisation of FtP data is analysed in the PSA (21) report concluded that the number and types (and level of details) of categories vary across regulators.
- For serious misconduct cases Bryce et al.(29) note inconsistencies in the structure of FtP procedures and approaches to investigation between UK regulators (including health and non-health).
- Inconsistencies in sexual misconduct investigations across three health care regulators observed in the report by Searle et al.(30)

These reports show that some inconsistencies exist across health and social care regulators. However, it is still not entirely clear what the extent of inconsistency is, and where and how it may affect decisions. FtP decisions are complex. As noted in Prof Sir William's report it:(1)

> "*.. is difficult to establish whether there is inconsistency in outcomes for what seem to be similar cases. Even in a single case where multiple professionals are involved, the actions and responsibilities of individual professionals will be different*".

There is also concern that inconsistency is greater at the earlier stages of the investigative process. The graph below uses the UK Health Regulator Comparative Data Report (28) produced by the GMC. It shows the number of complaints, investigated complaints, FtP hearings, and sanctions per 1,000 of registrants on each of the regulators databases. Whilst the functions of registrants are similar at the later part of the FtP process there appears more variation in the early phases of the process (see Figure 1).

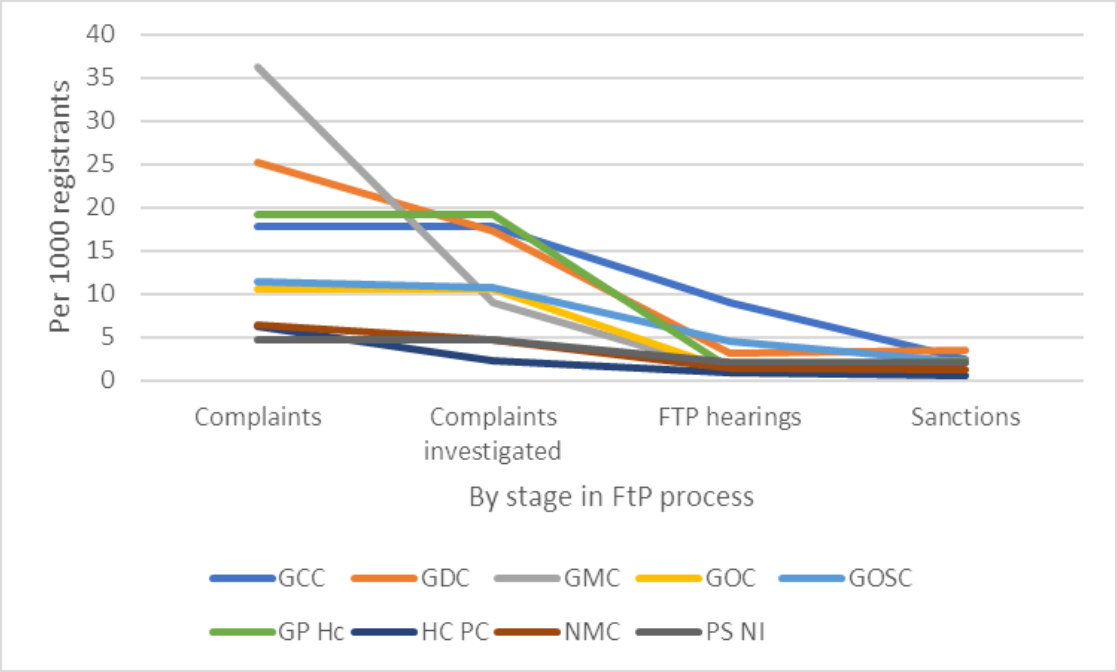

Figure 1: Variation in the proportion of registrants at the four stages of FtP investigation by regulator

In order to devise a methodology to inform future research into this complex and critical area it is vital to take into account the factors that influence consistency. These factors are detailed in the following section.

Factors influencing consistency

Factors that could influence the consistency of the investigative process have been grouped according to whether they operate at the wider societal or Macro level, the organisational or Meso level (in this case the regulator), or at the Micro level, the features pertaining to the registrant (see Figure 2).

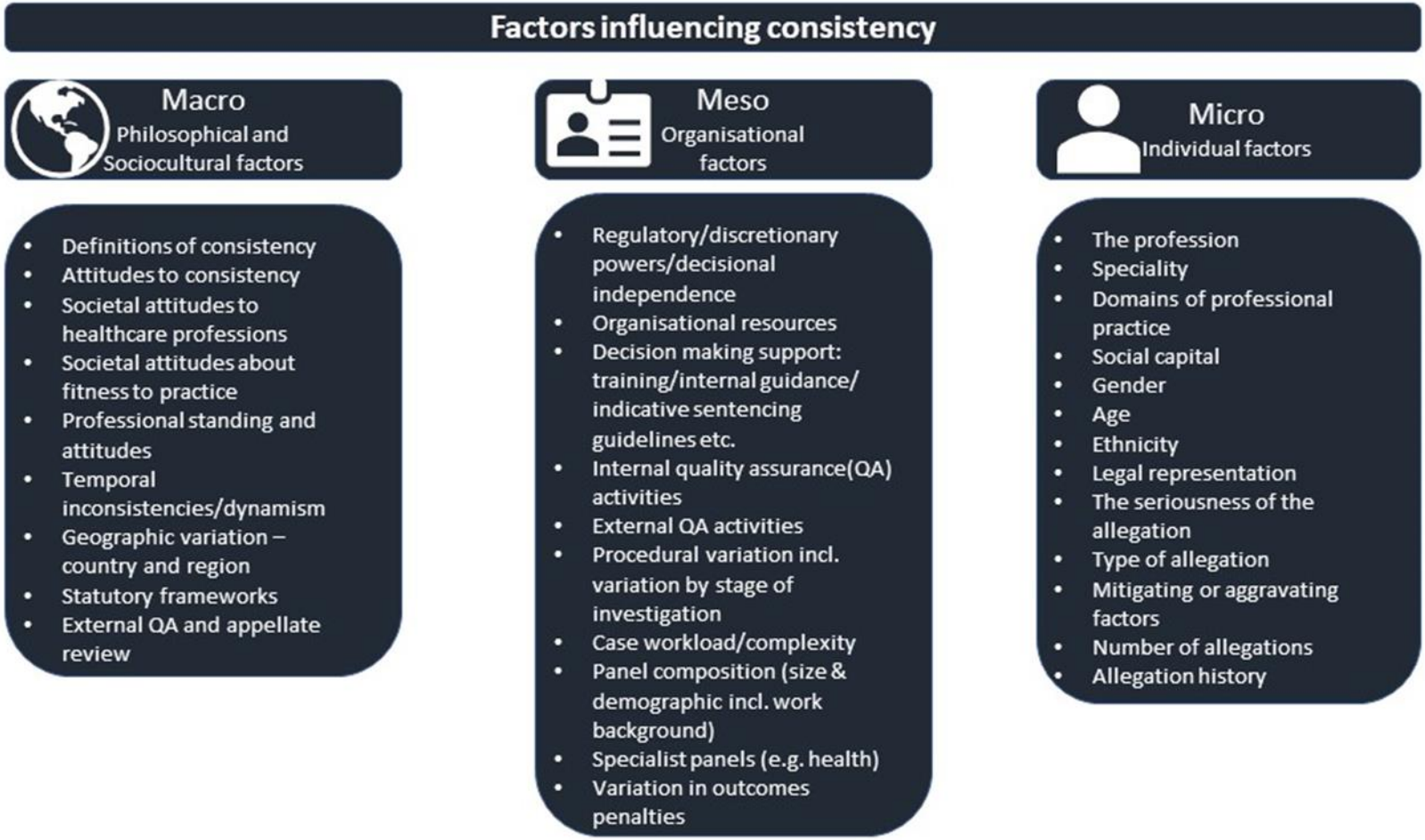

Figure 2: Factors influencing consistency

Macro level influences on consistency

There are a range of social, cultural and philosophical issues that could influence consistency. Firstly, how consistency is understood will influence both the public and professionals' attitudes towards it. Public protection has a common definition across all health and social care regulators and is enshrined in their legislation. It has three constructs: to protect patients, to maintain public confidence in the profession and to uphold standards. However, perceptions of what public confidence is and how regulators maintain it is poorly understood. We also do not know if societal attitudes towards professional misconduct in health and social care varies by profession or by offence.

> *"Little is understood about the type of behaviours and failings that might lead to the public losing confidence in the profession and therefore constitute grounds for regulatory action".* Williams Report (1)

Each professional group have their own codes of conduct, degrees of professional autonomy, and forms of practice and are bound by professional specific norms and values. Do these influence the professions' understanding of consistency? Do professional groups experience/enact different degrees of "cultural ambivalence" towards consistency and how does that impact on their professional discretion? How do professions perceive unprofessional behaviours and is there a spectrum of seriousness which influences consistency?

We know that the legal frameworks that the health and social care regulators work within vary, yet we do not have a detailed understanding about how these influence the consistency of the investigative process. We also know that geographic variation impacts (26) on consistency and that both societal and professional attitudes towards the various sorts of misconduct are dynamic and change over time. For example, sexual misconduct has been prevalent in the media (e.g. the #MeToo campaign) and this could influence future sentencing outcomes as this behaviour could now be considered far more unacceptable. Consistency in this way, may be a context specific phenomenon.

Meso level influences on consistency

We have already considered organisational issues impacting on consistency in the section on concerns, that is the various FtP processes each regulator has and the differences at the four stages of the investigative process, but there are other factors that may impact on consistency (see Figure 2). Organisational culture and philosophy will shape the ability of individuals involved in the investigative process to exhibit, or not, professional discretion and the ability to locally vary processes including decision-making. Organisation resources will have an impact too: under-resourced organisations may be deprived of the opportunity to invest in the practices and processes that provide the infrastructure to mitigate against inconsistency as well as to investigate its presence. Furthermore, caseload and complexity, especially in under-resourced organisations may negatively impact on consistent outcomes, as will a lack of training, guidance and internal quality assurance. The composition and nature of panels may also influence consistency.

Kevin Peel, dually registered with the NMC and HCPC (then HCPTS), illustrates how, despite judging the same offence, regulators reach differing outcomes.(31,32) Kevin Peel was caught trying to meet a teenage boy after engaging him in a sexually explicit conversation. For the same offence, Kevin was struck-off by the HCPC (then HCPTS) and suspended by the NMC. High profile cases involving co-offending individuals illustrate possible influences on outcome according to the degrees of professional discretion enacted by the various regulators. For example, Dr Bawa-Garba and nurse Amaro worked together at the time of the death of a six year old patient. Sanctions for these two health care professionals were different: Dr Bawa-Garba was suspended (33) while nurse Amaro was

struck off.(34) Similarly two health care professionals, a doctor and a physiotherapist, involved in the same dishonest situation of staging rugby players' injuries received different sanctions: Dr Wendy Chapman received a warning (35) and physiotherapist, Stephan Brennan, was struck-off.(36)

Micro level influences on consistency

There are three key areas of influence on consistency at the level of the individual professional: the professional background; individual demographics, and the nature of the case (including legal representation and multiple allegations). The 32 professions represented by the nine health and social care regulators have different roles and scopes of practice; each one affording variable interpretations and opportunities for professional misconduct. Even distinct professional groupings are not homogenous. For example, doctors specialise and work in diverse contexts, have differing clinical roles and patient demographics. There has been concern that some professions are treated more harshly than others; for example, the Bawa-Garba/Amaro case.(33,34)

Individual demographics including protected characteristics may influence consistency. We know from other research into differential attainment that ethnicity, gender and primary medical qualification have an impact on FtP outcomes and career progression. The Williams Report (1) also reported that:

> *"There was evidence to suggest that in both criminal and regulatory investigations there was a disproportionate number of Black, Asian and Minority Ethnic professionals involved."*

Searle et al. (30) identified that some cases appeared to be treated differently by different regulators. For example, qualifications fraud is related to failures to comply with GMC requirements and employer rules for doctors, while it is associated with police cautions and convictions in nurses and midwives. They also note the apparently different outcomes for sexual misconduct between regulators.

Implications for future research

Key methodological factors to consider for future research:

- The myriad factors that influence consistency at **the level of society, the organisation and the individual** (see Figure 2, p. 27);
- There was **variability between regulators** in terms of FtP processes including differences in investigative process at the four stages, legal frameworks, categorisation of allegations and outcomes, databases not readily available for statistical analysis and less detail for cases at earlier stages of the investigation;
- **Individual variability** in terms of the various professional standards, sociodemographic details, number of registrants, allegations categories and confounders for example multiple allegations;
- The need for comparative analysis across nine regulators with **variable number of registrants** in the FtP process.

Suggested methodological approach

Therefore, we argue that a holistic approach to the investigation of such a complex problem is necessary. Research needs to examine the Macro, Meso and Micro level influences in order to explore if inconsistency exists, if so who is affected, and crucially why; and to examine the causal factors which generate inconsistency. In this way any inconsistencies can be addressed. A robust and comprehensive programme of research needs to examine the regulatory contexts, procedural elements, individual registrant factors, and consistency as it relates both to outcome measures and decision-making.

We suggest analysing all regulators, a purposeful sample of allegation categories and consequent outcomes, and examining for inconsistency across all stages of the investigative process through rigorous mixed methods research methodologies and synthesising these findings to ensure robust and meaningful findings.

It is anticipated that this research could inform the development of future databases to ensure a method for ongoing evaluation of consistency after project closure. This could also be used to inform practice by for example developing agreed standards around consistency and recommend changes that ensure greater consistency.

Figure 3 provides an illustration of a suggested research framework and methodological approach by which consistency can be researched. Drawing upon our findings and research expertise we have designed a multilevel, multiphase, mixed methods approach as a methodology to investigate consistency. The framework would cover all nine healthcare regulators, take into account the key methodological considerations listed above and facilitate the development of a detailed comparative understanding of consistency.

We have broken down the research into a range of work packages – Consistency I through to Consistency VI. These appear in the middle of the figure and include:

Consistency I: Philosophical, legal and sociocultural influences on consistency;
Consistency II: Procedural influences at organisational level and understanding the causal factors for FtP outcomes;
Consistency III: Qualitative examination of variables influencing FtP outcomes;
Consistency IV: Quantitative examination of variables influencing FtP outcomes;
Consistency V: Evidence-based approach for ongoing evaluation of consistency;
Consistency VI: Legal representation.

The left-hand side of the figure identifies how the research will explore the Macro, Meso, and Micro factors known to influence consistency.

The right hand side of the figure shows the research questions pertaining to a future empirical study, and an indication about the nature of subsequent evidence that a study could generate, i.e. whether it will provide data on *what* is happening in regard to consistency but also where it will give insights about *why* it is occurring.

We will now describe our proposed methodological approach guided by the research questions below. For each question we provide a rationale, and strengths and limitations of our approach (Full study-Research question 3 is discussed in Scoping question 4, p. 41).

Full study-Research Question 1: What are the broad points of consistency and difference within the current fitness to practise procedures of the nine regulators?
Full study-Research Question 2: What is the potential impact of different variable factors within fitness to practise procedures?
Full study-Research Question 4: How might legal representation of registrants affect fitness to practise outcomes?

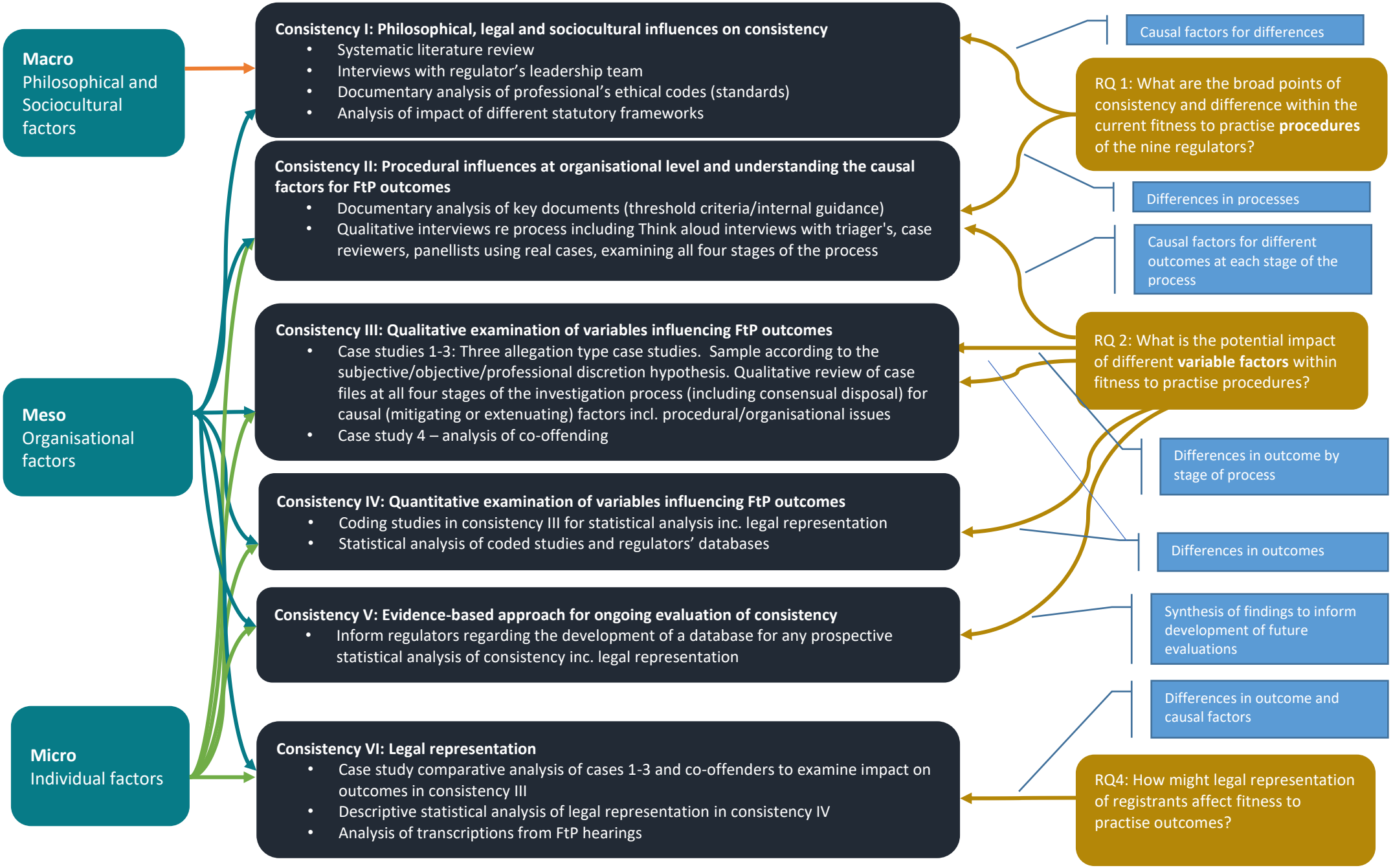

Figure 3: Research framework for examining consistency

Full study- RQ1: What are the broad points of consistency and difference within the current fitness to practise procedures of the nine regulators?

Consistency I: Philosophical, legal and sociocultural influences on consistency & Consistency II: Procedural influences at organisational level and understanding the causal factors for FtP outcomes

Systematic review

We recommend conducting a systematic literature review of academic and grey literature pertaining to consistency and difference within the current FtP procedures; as well as more generally to what inconsistency is and what factors influence it. Based on the findings of our scoping literature review (see above) we suggest reviewing literature beyond healthcare, and include other sectors, such as law**.**

Why this methodology is important to understand consistency:

A systematic review aims to provide a complete, exhaustive summary of current literature relevant to the research questions. Such a review will explore, critically appraise, and synthesise what is already known about (in)consistency in both healthcare regulation and decision-making more broadly. This is valuable to the research project as it will prevent duplication of work. It will also help to better understand what factors (and how) influence (in)consistency and identify knowledge gaps, and so will aid the formulation of novel research questions and, ultimately, result in meaningful and impactful findings. Information about Macro, Meso and Micro factors affecting consistency will be identified here.

*How and what needs to be done***:**

We suggest following established guidelines for systematic review (e.g. PRISMA)(37) to ensure transparency and thorough reporting of the results. Considering the broad topic, researchers may also want to consider using such software as EPPI-reviewer. To conduct this research, it will be necessary to have access to a variety of publication databases.

Interviews including "think aloud" interviews

We recommend conducting a) in-depth interviews with regulators' leadership teams (Consistency I), and b) a number of "think aloud" interviews with triagers/case assessors, hearing panels, and committee/panel members at each of the nine regulators (Consistency II).

a) Interviews

Why this methodology is important to understand consistency:

Of primary importance to understand consistency in the FtP processes within and across each of the nine regulators, is to collate**,** understand, and compare stakeholder perspectives and decision-making processes. Semi-structured, in-depth qualitative interviews are the most effective way to achieve this. This is because they offer the opportunity for decision makers to directly share their views on and experiences with (in)consistency, including understanding the impact of legal frameworks. Such interviews would enable researchers to gather experiences of, and beliefs about, consistency in decision-making in relation to regulation of FtP in healthcare. This style of data collection is valuable to the research project because it allows researchers the opportunity to interact with decision makers and ask *why* questions about the information that is shared. This approach to data collection will give insight into the Macro, Meso and Micro factors that affect consistency.

How and what needs to be done:
To conduct semi-structured in-depth interviews, researchers would need access to stakeholders who are willing to be interviewed. During in-depth interviews, participants would be asked questions relating to their opinions and experiences of (in)consistency in healthcare regulation. Conducting a discourse analysis (22) of stakeholder accounts would enable researchers to determine what explicit and implicit differences there might be across, and within, regulators' decision-making processes. Such an approach to analysis would be beneficial as it draws attention to the normative assumptions that decision makers are using during the FtP investigation process, and the impact that this has on the decisions that are subsequently made enabling comparison; across regulators and between decision makers within regulatory bodies. This would give insight into the Macro and Meso factors affecting consistency; but not forgetting that these shared experiences will also be first hand, descriptive accounts, and so will also give insight into the Micro factors that affect consistency.

b) "Think aloud" interviews.

Why this methodology is important to understand consistency:
We recommend conducting simultaneous "think-aloud" interviews (38,39) with FtP decision makers at each of the nine regulators at each of the four FtP stages. Conducting such interviews would enable researchers to determine what factors influence decision-making (and so the impact that such factors have on procedures and decisions) and what impact the presence/lack of variable factors has on decision-making outcomes. As well as offering insight into the potential impact of different factors, this method will offer insight into why such factors are influential.

How and what needs to be done:
To conduct these interviews, researchers would need access to triagers/case assessors, hearing panels, and committee/panel members at each of the nine regulators, who are willing to be interviewed. These interviews would involve presenting decision makers with real cases and asking them to narrate their "in the moment" decision-making processes. Attention could be paid to: what internal regulations are mentioned and drawn on to help with reviewing a case; what (if any) further evidence is requested to assist with making decisions; what "common sense" notions are drawn on to help make sense of the information in front of them (e.g. how and why they believe an accused behaviour is feasible); and what (if any) assumptions are made about unknown elements about the hypothetical case/registrant. These will be able to be compared both across regulators, and between FtP decision makers within regulatory bodies. These narrated actions will be first hand, descriptive accounts, and so will also give insight into the Macro, Meso and Micro factors that affect consistency.

Documentary analysis.

We recommend conducting a documentary analysis of each regulator's professional ethical codes or standards and the different statutory frameworks that they use to outline their minimum standards of practice (Consistency I), as well as the regulators' FtP policy documents and summary reports of investigations (Consistency II).

Why this methodology is important to understand consistency:
Analysis of this regulator-produced material will give insight into how the concept of "fitness to practise" is constructed and understood within each of these organisations, along with perceptions of the severity of FtP accusations. In other words, it enables an exploration of organisational-level attitudes and beliefs about FtP – thus aiding our exploration of consistency across regulators, by uncovering precisely what this concept is understood to be by each of them. This will enable a comparison across regulators and will highlight any inconsistencies in how this is understood to be.

This is valuable to the research because what FtP is thought to be will impact on how regulators investigate and assess it. If, for example, they perceive it to be a set of acquired skills and knowledge, FtP can be assessed by looking for a *lack* of these attributes. If, on the other hand, it is perceived to be the lack of a variety of negative traits, FtP can be assessed by looking for the presence of these traits.

How and what needs to be done:
We propose conducting a discourse analysis (22,40,41) of regulatory-produced materials. This approach to analysis essentially involves a process of looking at what "common sense" assumptions underpin the statements made within these documents, and exploring how and why they exist. Such an approach to analysis would be beneficial for two reasons. Firstly, it would permit a detailed exploration of whether, and in what ways, conceptualisations of FtP differ across regulators; and secondly, it would permit an explanation of how and why these conceptualisations are different. To conduct this analysis, researchers would need access to each of the regulators' FtP policy documents and instructions for assessors, investigators, and examiners; as well as regulators' professional ethical codes, the different statutory frameworks that they use to outline their minimum standards of practice, and any other relevant documentation related to FtP.

Strengths and limitations of the Consistency I and II methodologies:
Systematic review: The proposed systematic review will synthesise large bodies of research to give a better understanding of what is already known about (in)consistency in healthcare regulation, and also what can be learnt from other sectors. This is a powerful method for identifying existing knowledge if done properly. Systematic reviews, however, do have some limitations. For example, the process of synthesising research may lose the nuances evident in the literature reviewed, and focussing only on results means that the contextual factors of a study may be lost. In addition, all studies are awarded the same value if deemed relevant, reducing the weight of a particularly pertinent piece of work by mixing the relevant with the *not so* relevant. A systematic literature review is a time and cost-efficient way of exploring, identifying and understanding what is already known.

Interviews including "think aloud" interviews: Interviews are interactive and exploratory, and often provide incredibly useful insights by drawing out information from participants. Interviews are invaluable to understanding consistency in decision-making across the nine regulators, as it will demonstrate *real* "in the moment" (not assumed or described) decision-making processes. What transpires in these interviews can be compared within and across regulators, to develop a holistic understanding of how decisions are made and what variable factors impact on these processes.

Interviewing is time-consuming and requires resources at all stages of data collection and analysis. From identifying and approaching participants, to analysing transcripts, each related task requires dedicated researchers. Transcription of interview recordings is also costly.

Documentary analysis: discourse analysis of regulators' FtP documents offers novel insights and permits the development of a much more detailed understanding of the regulators themselves than, for example, a simple thematic analysis of their documents and reports. It will therefore enable a much more robust and substantial comparison of consistency across the nine regulators. The materials to analyse for this part of the project are mostly publicly available and so access will not be problematic.

Full study- RQ2: What is the potential impact of different variable factors within fitness to practise procedures?

Consistency III: Qualitative examination of variables influencing FtP outcomes

We recommend conducting a detailed analysis of case files, comprised of "real world" examples of regulators' FtP decision-making, including an examination of co-offending.

Why this methodology is important to understand consistency:
We recommend a detailed analysis of regulators' FtP case files across all four stages of the investigative process. These case files give of insight into how and why FtP investigations proceed and how decisions are made – in the sense that they reveal what information is requested, and used, by individuals/panels to make their decisions. This detailed analysis will highlight factors in the FtP process that impact on consistency. This approach allows both comparisons of consistency between the regulators but also the capacity to generate causal explanations if inconsistency is apparent.

How and what needs to be done:
We suggest a stratified case study approach. An informed hypothesis would be that some allegations would intrinsically have more opportunity for inconsistency. For example, some allegations are more "objective", i.e. criminal allegations which have evidence from the police/courts etc. These criminal cases give decision makers less opportunity for "professional discretion" and increase the likelihood of closely adhering to sentencing guidelines. A working hypothesis for these sorts of cases would be that there is a high degree of consistency across the nine regulatory bodies. There are also allegation categories where evidence is more likely to be subjective and those where professional discretion is more likely to come into play. Therefore, we suggest sampling case studies by allegation category using a sampling framework in Figure 4.

Figure 4: Hypothetical sampling strategy for Consistency III & IV

With the acknowledgement that FtP case files are typically unique to each particular investigation, and vary by regulator in terms of their size and level of detail, we propose conducting a content and discursive analysis of a sample of case files described above. This process would result in a detailed description of precisely what went on in relation to decision-making – e.g. *what* information was used to come to what decision (content analysis).(42,43) It would also result in a detailed description of *how* such information and evidence was used to come to that decision (discourse analysis) (22,40,41). This would permit an exploration of whether (i) there are differences both in what materials are

included across regulators and within regulators for the same/similar accusations; and (ii) whether there are noticeable patterns in or correlations between their content and FtP outcomes. We would suggest analysing a minimum of three files closed at each stage of the investigative process, for each of the three allegation case studies identified. Across all nine regulators this represents a total of 324 case files to analyse.

Strengths and limitations of this methodology:
Content analysis is a process of labelling (or coding) documents and typically offers simple descriptive data such as word frequencies and document lengths. Including a content analysis of case files will offer an overview of what information was used to come to what decision, by what regulators. Computer software can be used to aid this process and can greatly increase the number of texts that can be labelled. However, the reliability of using a computer programme to do so is questionable and the descriptive nature of this approach, whilst interesting, is not able to answer the "why" questions so needed in this research. The discourse analysis of case files will enhance the descriptions of them produced in the content analysis, enabling researchers to answer the "how" and "why" questions they need to. The proposed approach to the analysis of case files is time-consuming as it will require redaction of case files and the use of discourse analysis will require more resources than a simple content analysis of case files would. However, including a *discursive* exploration will offer novel insight and permits the development of a much more detailed understanding of the impact that variable factors have on the consistency of FtP decision-making. In addition to noting common themes and similarities in outcomes when a variable is present (e.g. "reflective statement"), it will also explain how and why these themes are common, as it will describe how and why a particular variable is able to have this impact.

Case study analysis of co-offenders to examine impact on outcomes.
We also suggest a fourth case study looking at a sample of cases with co-offenders. This will also enable researchers to draw out in what way regulators are consistent and where they vary. We recommend conducting a comparative analysis of cases, in which registrants from different regulators were involved in the same incident, to examine any differences in outcomes and in registrant representation.

Why this methodology is important to understand consistency:
A detailed analysis of case files is recommended to examine consistency across regulators, using a similar approach to the case file data described above. Cases could be used to examine process and outcomes in cases where more than one registrant was involved in the same incident. A similar thematic and discourse analysis could be conducted to see how the same incident is handled and how decisions are made by different regulators.

How and what needs to be done:
As above, a sample of case files in which the same incident has been investigated by different regulators for different registrants would be needed for an in-depth qualitative analysis. Thematic analysis could be used to determine the similarities and differences between the various regulators' evidence-gathering and processes, and discourse analysis could be used to examine how the regulators made their decisions, and whether these differ among the regulators.

Strengths and limitations of this methodology:
This level of analysis is time consuming but will provide rich insight into the decision-making processes of regulators, not only in relation to similar allegations, but directly comparable within the same incidents. As there may be differences in representation across the affected registrants, this type of comparative analysis will also highlight any significant issues in this area.

Consistency IV: Quantitative examination of variables influencing FtP outcomes

To gain a quantitative understanding about consistency we suggest combining regulators' databases (over the last three years) for statistical analysis. Furthermore, we suggest these should be supplemented by coding a sample of 1,300 case studies for the additional detailed contextual and demographic data that they contain. We recommend using the same case study approach described in Consistency III for quantitative examination of the variables influencing FtP outcomes.

Why this methodology is important to understand consistency:
The existing FtP databases described in section two of this report hold valuable information and are large datasets, but alone are not adequate for the quantitative assessment of consistency in FtP. They do not currently reliably record all the variables which need to be examined in a study of consistency. FtP databases need to be merged and important additional variables entered manually from the case files (e.g. legal representation, consideration of mitigation circumstances). Finally, the newly created research database needs to be anonymised for research purposes.

How and what needs to be done:
Based on the secondary data in the GMC UK Health Regulator Comparative Data Report (28) (see Table 11), and taking into account how many complaints each regulator investigates, we estimate that the minimum sample size needed for Consistency IV is 1,300.

Table 11. Comparisons of FtP outcomes among the nine health and social care regulators (2015/2016)

Regulators	Investigated complaints	Warnings	Warnings per 100	Conditions/ undertaking	Conditions/ undertakings per 100	Suspension	Suspension per 100	Erasure	Erasure per 100	Sanction per 100
GCC	50	3	**6.0**	2	**4.0**	0	**0.0**	2	**4.0**	**14**
GDC	1870	187	**10.0**	78	**4.2**	77	**4.1**	37	**2.0**	**20**
GMC	2090	148	**7.1**	168	**8.0**	93	**4.4**	78	**3.7**	**23**
GOC	223	28	**12.6**	3	**1.3**	4	**1.8**	9	**4.0**	**20**
GOsC	49	3	**6.1**	1	**2.0**	3	**6.1**	3	**6.1**	**20**
GPhC	1437	83	**5.8**	11	**0.8**	47	**3.3**	28	**1.9**	**12**
HCPC	787	29	**3.7**	37	**4.7**	60	**7.6**	71	**9.0**	**25**
NMC	3245	119	**3.7**	152	**4.7**	277	**8.5**	261	**8.0**	**25**
PSNI	11	2	**18.2**	0	**0.0**	0	**0.0**	3	**27.3**	**45**
TOTAL		602		452		561		492		

Note. Based on data from the UK Health Regulator Comparative Data Report (28)

This methodology will create a database with a comprehensive range of variables for assessing the consistency of FtP outcomes. Importantly it can more fully assess the impact of the Micro level

variables, listed in Figure 2, p. 27, on consistency as these can be coded from the 1,300 case files and added to the database. For example, additional elements that could be included are:

- Individual demographics – ethnicity, place of primary qualification, age, etc.;
- Engagement with the process and attendance at hearings;
- Contextual matters in the case;
- Legal representation.

In addition, this methodology will help to explore the impact of multiple allegations (e.g. number of allegations, different types of allegations) on outcome (and outcome severity). Depending on the completeness of regulators' databases at each of the four stages of the investigative process, a further analysis could be performed to examine consistency at the various stages.

Strengths and limitations of this methodology:
This statistical analysis will help to better understand which factors link to FtP outcomes. Coding cases will give a unique opportunity to quantify some of the process factors which are rarely available for statistical explorations (e.g. contextual matters). This is important as from previous studies it is known that some groups of health care professionals are more severely sanctioned (e.g. men; see Scoping question 1), however the reasons for such observations are not always clear. This proposed quantitative analysis will help 1) to test otherwise rarely explored factors which may have an impact on outcomes; 2) to combine various factors together to help explain some of the previous observations (e.g. do men have multiple allegations more often and is this why they are more severely sanctioned); and 3) to demonstrate consistency across regulators.

The main limitation of this approach is that it will be time consuming, including a process to ensure the accuracy of the data (double entry). The very small numbers of cases that some of the regulators have will be problematic and unlikely to present statistically significant results; however, this is in part mitigated against by Consistency III.

Consistency V: Evidence-based approach for ongoing evaluation of consistency

This research is likely to identify factors that will impact on consistency. Some factors may be more influential than others and the research will be able to inform regulators about what data is the most critical to record. This could then be used by regulators in the ongoing evaluation of consistency in FtP decisions. Furthermore, research of this nature could be used to improve practice.

Full study- RQ4: How might legal representation of registrants might affect fitness to practise outcomes?

Exploring how legal representation may impact on the outcomes of the FtP process has already been detailed in the response to research question two, in particular in the sections on Consistency III and V. Here we suggest using case files and statistical analysis to explore the impact of legal representation qualitatively and quantitatively, through a descriptive statistical analysis of legal representation, and a case study comparative analysis of cases 1-3 and co-offenders to examine impact on outcomes.

To understand in greater detail how legal representation affect FtP outcomes we suggest an additional methodological approach.

Analysis of hearings.
We recommend conducting a detailed analysis of hearing transcripts.

Why this methodology is important to understand consistency:
There are several ways that a hearing transcript could be approached to examine the effect of registrant representation. As hearing transcripts provide a seemingly accurate account of what was discussed in the hearings, this data could be used for qualitative analysis. Not only could this be used to compare the proceedings across regulators, but it could also be used to examine how the hearings compare when the registrant does or does not have some kind of representation. Firstly, it will give an indication of how much talk is produced by whom: how much does the registrant representative contribute to the discussions, and does this differ when the registrant represents themselves? Secondly, the transcripts can be used to investigate how a hearing progresses when the registrant has representation, compared with when the registrant has no representation, or is absent from the proceedings. For example, how much time is devoted to the registrant's perspective with and without representation, and in their presence or absence?

How and what needs to be done:
In order to answer these questions, discourse analysis (22) could be used to look at the language used, and the discourses that emerge during hearings; this is particularly useful for taking into account any issues around power or fairness. The analysis could also be informed by conversation analysis,(23,44) which is a way of examining how interaction works at a micro-level of detail in various settings.

This would include looking at: how *sequences of talk* progress, for example whether and how particular questions are responded to by the various participants in the hearing; *turn-taking*, for example how speakers are chosen or choose to speak themselves during the hearing; the use of language to repair, which concerns if and how speakers seek to correct misunderstandings that might occur in the hearings and if/how they are corrected; and the terminology that speakers use in their talk. As a hearing is a high-stakes scenario, where a registrant could have serious sanctions imposed on their practice and therefore livelihood, looking at these interactions in this level of detail could illuminate some key yet under-the-radar elements of the process.

To carry out this kind of analysis it is necessary to have access to cases containing complete accurate transcripts of the case hearings. For some of the cases in the S29 database, transcripts of the hearings from across the regulators have been included in the case documentation. A sample of fifty transcripts would allow for an in-depth examination of the interactional detail of hearings; smaller numbers are generally sufficient for this type of qualitative analysis, but the final sample size would be determined in consultation with qualitative sampling theories and models.(45,46) This sample of fifty should contain equal numbers of hearings in which registrant representation was present, and hearings in which there was no representation. If possible, and dependent on data availability in the database and from the regulators, these sub-sets could be broken down yet further to include: 10-15 with representation and the registrant present; 10-15 with representation but the registrant absent; 10-15 with no representation but the registrant present; and 10-15 with no representation and the registrant absent. This would allow a detailed analysis of the various levels of registrant engagement with the hearing process.

Strengths and limitations of this methodology:
Working through the hearing transcripts in this way has the potential to unearth interesting findings at both a broad and narrow level. By analysing the discourses running through the talk recorded in the

transcripts, any broad themes to do with power, fairness, and the organisations involved will emerge. By looking in detail at the talk produced by the various speakers it will be possible to see how issues emerge at a more individual or interpersonal level. By examining the hearing transcripts at both of these levels, it should be possible to determine how the presence or absence of representation, and of the registrant in the proceedings, affects how the hearings progress, and whether this has any bearing on the fairness of the process. However, there is a cost implication, as transcription of full hearings, which can take place over several days, is expensive. While some transcripts are included in the S29 database, only a relatively small number of cases include registrants who have been legally represented (10 cases). Therefore, additional transcripts will need to be requested.

Scoping question 4. What is the scope of a potential full research project, including costings, timelines, etc.?

Full study- RQ3: What is the potential for different academic disciplines to contribute to addressing the question of consistency, in light of the range of variable factors?

Considering the complexity of the topic, we recommend that in order to investigate (in)consistency in FtP it is essential to employ a mixed methods approach. To achieve this, a multidisciplinary team with broad areas of expertise should be involved in undertaking such a project. First, a researcher with experience in systematic reviews (meta-analysis) should be involved in this project to be able to answer the first bullet point in Consistency I (see Figure 3). Second, we suggest involving researchers with an expertise in qualitative analysis to answer research questions on Consistency I, II, III, V, and VI. These researchers should be experienced in documentary analysis (e.g. policy, guidelines), conducting interviews including think-aloud interviews, ethnographic studies, and ideally having knowledge in various qualitative analysis such as thematic analysis and discourse analysis. Such specialists might come from various academic backgrounds, e.g. education, sociology, psychology, or linguistics. Third, researchers who are experienced in analysing quantitative data will be needed to investigate Consistency IV, V and VI. Such researchers should have experience in various statistical methodologies (e.g. chi-square, logistic regression), but also have experience in dealing with secondary data (not collected for research purposes) and ideally have expertise in analysing FtP or other types of court decisions. This suggestion is based on our observation of the complexity of the data on FtP and the unique challenges related to datasets like this. Quantitative researchers may also be from various academic backgrounds, e.g. statistics, epidemiology, psychology. Fourth, a data cleaner with attention to detail could be responsible for data extraction from the case files (Consistency IV) and a research assistant as well as administrator to support researchers throughout this study is essential considering the large number of administrative and research organisation tasks. We also want to highlight that the research team should seek advisors from the legal sector and from those who are closely involved in FtP procedures.

Costs

Our estimated detailed costings have been removed from this publicly-available report, as it is deemed commercially sensitive.

Timeline

This research is complex and readily available sources of data to analyse in exploring consistency are currently unavailable. Therefore, we suggest that a future research project of this nature would take approximately three years to complete. Our timeline is detailed in the Gantt chart below takes into account negotiations and agreements regarding data sharing that will be crucial to the success of the project (a more easily readable Gantt chart is attached as a separate file).

GANTT CHART FOR CONSISTENCY RESEARCH SEPT 2019 - SEPT 2022

Project Management	**Before start**	**Sep-19**	**Oct-19**	**Nov-19**	**Dec-19**	**Jan-20**	**Feb-20**	**Mar-20**	**Apr-20**	**5/1/2020**	**Jun-20**	**Jul-20**	**Aug-20**	**Sep-20**	**Oct-20**	**Nov-20**	**Dec-20**	**Jan-21**	**Feb-21**	**Mar-21**	**Apr-21**	**May-21**	**Jun-21**	**Jul-21**	**Augs-21**	**Sep-21**	**Nov-21**	**Jan-22**	**Feb-22**	**Mar-22**	**Apr-22**	**May-22**	**Jun-22**	**Jul-22**	**Aug-22**	**Sep-22**
Stakeholder meetings	Indentify, recruit, TORs	x				x				x				x				x				x				x					x	x			
Cross-project strands meeting	Review study design	x		x			x		x		x		x		x		x		x		x		x		x		x		x			x	x	x	x
Ethics and data protection	Apply Ethics/R&D	x	x	x	x																														
Data sharing agreements	x	x	x	x	x																														
Governance and reporting	x												Progress report											Progress report							Write up final report			**Submit final report**
Recruit staff	x	x	x	x																														
Project Finance	Set up systems	Review cost data systems	finance report									finance report									finance report					finance report									
Dissemination																																			Papers and presentations
Consistency I: Philosophical, legal and sociocultural influences on consistency																																			
Systematic review			x	x	x	x	x	x	x																										
Interverviews with regulator's leadership team								x	x	x	x	x	x	x																					
Documentary analysis				x	x	x																													
Analysis of statutory frameworks			x	x																															
Consistency II: Procedural influences at organisational level and understanding the causal factors for FtP outcomes																																			
Interviews and think aloud interviews with case assessors/FtP teams/ panels/legal assessors													x	x	x	x	x	x	x																
Documentary analysis									x	x	x	x																							
Consistency III: Qualitative examination of variables influencing FtP outcomes																																			
Case studies 1-3						redaction	redaction	redaction	redaction	redaction	redaction	redaction	redaction	redaction	redaction	redaction	redaction	x	x	x	x	x	x	x	x	x	x	x	x						
Case studies 4																		x	x	x	x	x	x	x	x	x	x	x							
Consistency IV: Quantitative examination of variables influencing FtP outcomes																																			
Analysis of cases					redaction	redaction	redaction	Coding	Coding	Coding	Coding	Coding	Coding	Coding	Coding	Coding	Coding	x	x	x	x	x	x	x	x	x	x								
Analysis of regulators databases								x	x	x	x	x	x	x	x	x	x	x	x																
Consistency V: Evidence-based approach for ongoing evaluation of consistency																														x	x	x	x		
Consistency VI: Legal representation																																			
Analysis of hearings																		x	x	x	x	x	x	x	x	x	x	x							
Project synthesis																													x	x	x	x	x		

Figure 5: Project Gantt chart (please see separate excel file)

References

1. Williams PSN. Gross negligence manslaughter in healthcare. 2018;(June). Available from: www.nationalarchives.gov.uk/doc/open-government-licence/
2. Singh P, Mizrahi E, Korb S. A five-year review of cases appearing before the General Dental Council's Professional Conduct Committee. Br Dent J. 2009;206(4):217–23.
3. Yates J, James D. Risk factors at medical school for subsequent professional. BMJ. 2010;340(c2040).
4. Humphrey C, Hickman S, Gulliford MC. Place of medical qualification and outcomes of UK General Medical Council "fitness to practise" process: Cohort study. Bmj. 2011;342(7802):1–10.
5. Unsworth J. Student professional suitability: Lessons from how the regulator handles fitness to practise cases. Nurse Educ Today [Internet]. 2011;31(5):466–71. Available from: http://dx.doi.org/10.1016/j.nedt.2010.08.006
6. Campbell G, Rollin AM, Smith AF. Cases relating to anaesthetists handled by the UK General Medical Council in 2009: Methodological approach and patterns of referral. Anaesthesia. 2013;68(5):453–60.
7. Unwin E, Woolf K, Wadlow C, Dacre J. Disciplined doctors: Does the sex of a doctor matter? A cross-sectional study examining the association between a doctor's sex and receiving sanctions against their medical registration. BMJ Open. 2014;4(8):1–9.
8. Donaldson LJ, Panesar SS, McAvoy PA, Scarrott DM. Identification of poor performance in a national medical workforce over 11 years: An observational study. BMJ Qual Saf. 2014;23(2):147–52.
9. Bourne T, Wynants L, Peters M, Van Audenhove C, Timmerman D, Van Calster B, et al. The impact of complaints procedures on the welfare, health and clinical practise of 7926 doctors in the UK: A cross-sectional survey. BMJ Open. 2015;5(1).
10. Unwin E, Woolf K, Wadlow C, Potts HWW, Dacre J. Sex differences in medico-legal action against doctors: A systematic review and meta-analysis. BMC Med [Internet]. 2015;13(1). Available from: http://dx.doi.org/10.1186/s12916-015-0413-5
11. Gallagher CT, Greenland VAM, Hickman AC. Eram, ergo sum? A 1-year retrospective study of General Pharmaceutical Council fitness to practise hearings. Int J Pharm Pract. 2015;23(3):205–11.
12. Gallagher CT, De Souza AI. A retrospective analysis of the GDC's performance against its newly-approved fitness to practise guidance. Br Dent J [Internet]. 2015;219(5):E5. Available from: http://dx.doi.org/10.1038/sj.bdj.2015.686
13. Bourne T, Vanderhaegen J, Vranken R, Wynants L, De Cock B, Peters M, et al. Doctors' experiences and their perception of the most stressful aspects of complaints processes in the UK: An analysis of qualitative survey data. BMJ Open. 2016;6(7):1–10.
14. Brindley J. Reflection on fitness to practise. Br Dent J [Internet]. 2016;221(8):495–8. Available from: http://dx.doi.org/10.1038/sj.bdj.2016.779
15. Bourne T, De Cock B, Wynants L, Peters M, Van Audenhove C, Timmerman D, et al. Doctors' perception of support and the processes involved in complaints investigations and how these relate to welfare and defensive practice: a cross-sectional survey of the UK physicians. BMJ Open. 2017;7(11):e017856.
16. Tiffin PA, Paton LW, Mwandigha LM, McLachlan JC, Illing J. Predicting fitness to practise events in international medical graduates who registered as UK doctors via the Professional and Linguistic Assessments Board (PLAB) system: a national cohort study. BMC Med [Internet]. 2017;15(1):1–16. Available from: http://dx.doi.org/10.1186/s12916-017-0829-1
17. Mehdizadeh L, Potts HWW, Sturrock A, Dacre J. Prevalence of GMC performance assessments in the United Kingdom: a retrospective cohort analysis by country of medical qualification. BMC Med Educ. 2017;17(1):1–13.

18. Neville P. Social media and professionalism: A retrospective content analysis of Fitness to Practise cases heard by the GDC concerning social media complaints. Br Dent J [Internet]. 2017;223(5):353–7. Available from: http://dx.doi.org/10.1038/sj.bdj.2017.765
19. Leigh J, Worsley A, McLaughlin K. An analysis of HCPC fitness to practise hearings: Fit to Practise or Fit for Purpose? Ethics Soc Welf. 2017;11(4):382–96.
20. Wakeford R, Ludka K, Woolf K, Mcmanus IC. Fitness to practise sanctions in UK doctors are predicted by poor performance at MRCGP and MRCP (UK) assessments: data linkage study. BMC Med. 2018;16(230):1–17.
21. Professional Standards Authority. Categorisation of fitness to practise data: a description of UK health and care professional regulators' categorisation of fitness to practise allegations. 2017;(December). Available from: https://www.professionalstandards.org.uk/publications/detail/categorisation-of-fitness-to-practise-data
22. Jørgensen M, Phillips L. Discourse Analysis as Theory and Method. London: SAGE Publications; 2002.
23. Heritage J. Conversation analysis and institutional talk: analysing data. In: Silverman D, editor. Qualitative research: Theory, method and practice. London: Sage Publications; 1997. p. 161–82.
24. Sentencing Council. Analytical Note : the Resource Effects of increased consistency in sentencing [Internet]. 2011. Available from: https://www.sentencingcouncil.org.uk/publications/item/consistency-in-sentencing/
25. Krasnostein S, Freiberg A. Pursuing consistency in an individualistic sentencing framework: If you know where you're going, how do you know when you've got there? Law Contemp Probl [P] [Internet]. 2013;76(1):265–88. Available from: http://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=4354&context=lcp
26. Legomsky SH. Learning to Live with Unequal Justice: Asylum and the Limits to Consistency. Stanford Law Rev [Internet]. 2008 [cited 2019 Apr 11];60(2):413–74. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1085265
27. Professional Standards Authority. Right-touch reform. 2017.
28. General Medical Council for Chief Executive Steering Group. UK health regulator comparative data report 2016 [Internet]. 2016. Available from: http://www.gmc-uk.org/static/documents/content/UK_health_regulator_comparative_report_FINAL_220217.pdf
29. Bryce M, Archer J, Brennan N, Burns L, Brien TO, Price T. Fitness to practise : Impairment and Serious Misconduct A narrative synthesis review. 2018;(March).
30. Searle RH, Rice C, Mcconnell AA, Dawson JF. Bad apples? Bad barrels? Or bad cellars? Antecedents and processes of professional misconduct in UK Health and Social Care: Insights into sexual misconduct and dishonesty. Psa. 2017;
31. Nursing and Midwifery Council. Fitness to Practise Committee Substantive Hearing. Kevin Peel case. [Internet]. Available from: https://www.nmc.org.uk/globalassets/sitedocuments/ftpoutcomes/2018/april-2018/reasons-peel-ftpcsh-57782-20180406.pdf
32. Health and Care Professions Tribunal Service. Conduct and Competence Committee Panel. Kevin Peel Case [Internet]. 2018. Available from: https://www.nmc.org.uk/globalassets/sitedocuments/ftpoutcomes/2016/september/reasons-cafferkey-cccsh-49242-20160914.pdf?_t_id=1B2M2Y8AsgTpgAmY7PhCfg%3D%3D&_t_q=pauline+cafferkey&_t_tags=language%3Aen%2Csiteid%3Ad6891695-0234-463b-bf74-1bfb02644b38&_t_ip=8
33. Medical Practitioners Tribunal Service. Record of Determination. Dr Hadiza Bawa-Garba Case. [Internet]. 2018. Available from: https://www.mpts-uk.org/-/media/mpts-rod-files/dr-hadiza-bawa-garba-19-december-2018.pdf

34. Nursing and Midwifery Council. Conduct and Competence Committee Substantive Hearing. Isabel Maria Amaro Case. [Internet]. 2016. Available from: https://www.nmc.org.uk/globalassets/sitedocuments/ftpoutcomes/2016/august/reasons-amaro-cccsh-36269-20160804.pdf
35. General Medical Council. Fitness to Practise Panel. Dr Wendy Chapman Case. 2010.
36. Health and Care Professions Tribunal Service. Conduct and Competence Committee Panel. Stephen Brennan Case. 2010.
37. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339.
38. Fonteyn M, Kuipers B, Grobe S. A Description of Think Aloud Method and Protocol Analysis. Qual Health Res. 1993;3(4):430–41.
39. Charters E. The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods. Brock Educ. 2003;12(2):68–82.
40. Glynos J, Howarth D, Norval A, Speed E. Discourse Analysis: varieties and methods. ESRC Natl Cent Res Methods [Internet]. 2009;14(August):1–41. Available from: http://eprints.ncrm.ac.uk/796/1/discourse_analysis_NCRM_014.pdf%5Cnhttp://eprints.ncrm.ac.uk/796/
41. Foucault M. The order of dicsourse. In: Young R, editor. Untying the Text: A Post-Structuralist Reader [Internet]. Boston: Routledge & Kegan Paul; 1981. p. 51–78. Available from: http://eprints.ncrm.ac.uk/796/1/discourse_analysis_NCRM_014.pdf%5Cnhttp://eprints.ncrm.ac.uk/796/
42. Hsieh H-F, Shannon SE. Three Approaches to Qualitative Content Analysis. Qual Health Res [Internet]. 2005 Nov 1 [cited 2019 Apr 11];15(9):1277–88. Available from: http://journals.sagepub.com/doi/10.1177/1049732305276687
43. Downe-Wamboldt B. Content analysis: Method, applications, and issues. Health Care Women Int [Internet]. 1992 [cited 2019 Apr 11];13:313–21. Available from: https://www.tandfonline.com/action/journalInformation?journalCode=uhcw20
44. Sidnell J, Stivers T. The Handbook of Conversation Analysis [Internet]. Sidnell J, Stivers T, editors. Chichester, UK: John Wiley & Sons, Ltd; 2012 [cited 2019 Apr 5]. Available from: http://doi.wiley.com/10.1002/9781118325001
45. Kirsti Malterud, Volkert Dirk Siersma, Ann Dorrit Guassora. Sample Size in Qualitative Interview Studies: Guided by Information Power. Innov Methods [Internet]. 2016 [cited 2019 Apr 5];26(13):1753–60. Available from: https://journals.sagepub.com/doi/pdf/10.1177/1049732315617444
46. Saumure K, Given LM. Data Saturation In: The SAGE Encyclopedia of Qualitative Research Methods. 2008 [cited 2019 Apr 5]; Available from: http://dx.doi.org/10.4135/9781412963909.n99

Appendix A.

CHECK LIST FOR DISCUSSION

GENERAL QUESTIONS

Are these the correct allegation categories used for fitness to practise (FtP) complaints? If not, what allegations do you consider for investigation and how do you categorise these allegations?

	Allegations
GMC	https://www.gmc-uk.org/-/media/documents/dc4596-ce-decision-guidance---annex-g---examples-of-failures-to-meet-standards-62041464.pdf
NMC	https://www.nmc.org.uk/ftp-library/understanding-fitness-to-practise/fitness-to-practise-allegations/
HCPC	https://www.hcpc-uk.org/concerns/what-we-investigate/fitness-to-practise/ [types of cases]
GPhC	https://www.pharmacyregulation.org/sites/default/files/document/good_decision_making_investigations_and_threshold_criteria_guidance_january_2018.pdf [p 8; 2.7]
GDC	https://www.gdc-uk.org/professionals/ftp-prof
GOC	https://www.optical.org/en/Investigating_complaints/fitness-to-practise-guidance/index.cfm [Guidance on acceptance criteria; 1.9]
GOsC	file://ad.ucl.ac.uk/homee/rmheame/Documents/5ic-decision-making-guidance-2018.pdf [p.14]
GCC	https://www.gcc-uk.org/concerns/raising-a-concern-about-a-fellow-chiropractor/ [Concerns you should raise with us]
PSNI	http://www.psni.org.uk/psni/fitness-to-practise/ [Impairment to fitness to practise]

Is this a correct list of possible FtP outcomes?

	Case examiners	**Hearing**
GMC	https://www.gmc-uk.org/concerns/information-for-doctors-under-investigation/how-we-investigate-concerns/making-our-decision	https://www.gmc-uk.org/concerns/information-for-doctors-under-investigation/our-sanctions/referral-for-a-hearing-at-the-mpts
NMC	https://www.nmc.org.uk/concerns-nurses-midwives/support-for-patients-families-and-public/how-we-reach-an-outcome/	https://www.nmc.org.uk/concerns-nurses-midwives/hearings/our-panels-case-examiners/fitness-to-practise-committee/
HCPC	https://www.hcpc-uk.org/concerns/how-we-investigate/the-investigating-committee/	https://www.hcpts-uk.org/globalassets/hcpts-site/publications/policy/indicative-sanctions-policy.pdf [7]
GPhC	https://www.pharmacyregulation.org/sites/default/files/decision-making_at_an_ftpc.pdf	https://www.pharmacyregulation.org/sites/default/files/decision-making_at_an_ftpc.pdf
GDC		https://www.gdc-uk.org/api/files/Guidance%20for%20the%20Practice%20Committees%20-%20Indicative%20Sanctions%20Guidance.pdf [7.4]
GOC	https://www.optical.org/download.cfm?docid=5E42702F-B167-4321-82B19EBA232C7C6A [36]	https://www.optical.org/download.cfm?docid=4E0A56F2-08BA-4AB3-BAD9744B5DA591A5 [29]
GOsC	file://ad.ucl.ac.uk/homee/rmheame/Documents/5ic-decision-making-guidance-2018.pdf	https://www.osteopathy.org.uk/news-and-resources/document-library/fitness-to-

		practise/hearings-and-sanctions-guidance/ [11]
GCC	https://www.gcc-uk.org/concerns/the-investigation-process/	https://www.gcc-uk.org/UserFiles/Docs/Guidance%20on%20Sanctions%20April%202018.pdf [B3.78]
PSNI	http://www.psni.org.uk/wp-content/uploads/2012/12/Sanctions-available-to-fitness-to-practise-committees.pdf [Sanctions available to the Pharmaceutical Society NI's Scrutiny Committee]	http://www.psni.org.uk/wp-content/uploads/2012/12/Sanctions-available-to-fitness-to-practise-committees.pdf [Sanctions available to the Pharmaceutical Society NI's Statutory Committee]

What data sources do you hold on cases at each stage of the FtP process?

- Initial complaint/enquiry/concern
- Triage
- Investigation
- Hearing/prosecution

Do you record what data is available? How?

Generally speaking, how long do you hold on to all the data regarding a fitness to practise concern?

STATISTICAL DATABASES

Do you have a statistical database for the fitness to practise process?

If so, what is the size of your database, i.e. how many healthcare professionals/complaints do you have data for?

How many years' worth of data do you have?

Do you record how long case was open?

What information does that database include for each stage of the process?

- Initial complaint/enquiry/concern
- Triage
- Investigation
- Hearing/prosecution

How data is recorded, e.g. does every row of the dataset represent complaint or health care professional?

What variables are recorded in the dataset? *(e.g. demographic characteristics of registrant: gender ethnicity age, enquiry source, etc.)*

Has there been any important changes made during this period in recording data as well as categorisation of complaints/changes in processes, etc.?

Is there any information about decision-making*? (e.g. examiners notes; exam scores, etc.)*

Do you include data on legal representation in the statistical database? If so, in what detail.

- Who is providing legal representation
- The stage of the process where representation starts

Do you list all complaints/enquiries against the subsequent outcomes?

CASE FILES/BUNDLES

How many case files do you hold?

What is an average length of case files? In your opinion, does it vary by allegation category and how?

Which parts of the FtP process do you open/hold case files for?

- Initial complaint/enquiry/concern
- Triage
- Investigation
- Hearing/prosecution
- Review or consensual panels

From our experience we think that there is no common format in which case bundles are curated/structured. Is that correct?

What content is typically included in a case file?

  - Do they have identifiable details of the registrant and/or other parties?
  - What demographic/contextual data do you collect about the registrants undergoing fitness to practise? *(e.g. gender, ethnicity)*

Do you record if they have had legal representation or not? If so, in what detail.

- Who is providing legal representation
- The stage of the process where representation starts

Who makes a decision about what information is included in cases files?

FITNESS TO PRACTISE HEARINGS

Where do you hold fitness to practise hearings?

We understand that there are full transcripts of panels/hearings? How many transcripts of hearings do you hold?

Are determinations publicly or/and privately available? What information is available in publicly/privately available determinations?

Do you record any other information about the hearings other than that that would appear in the registrants case file?

Could you tell us a bit more about the composition of the panel members:

- Are there any rules guiding panel selection *(e.g. professional experience; diversity - requirement for 1 woman/BME representative on them)*;

- Do they have a legal chair?
- Do they have a legal assessor?

OTHER DATA
Is there any other data that you are aware of that may be helpful in any future research? For example, any internal quality assurance reviews or other research?

What data do you hold on review/consensual panels?

Do you have internal guidance for decision-making which is not publicly available? If so, describe what.

DATA SHARING
Can researchers get access to the database/s, casefiles, non-publicly available guidance or transcripts?

What would the process be to get permissions?

Could the statistical database be easily amended and anonymised so that the research team could see its structure?

Is this the same for all cases i.e. those cases closed early in the investigation?

Do you have identifiable details of the registrant and/or other parties redacted? If yes, can researchers have access to redacted data?

For a bigger research project which would likely to require redaction what issues would that raise for you?

Is there anything else that important to consider that we haven't asked?