Runners, repeaters, strangers and aliens: operationalising efficient output disclosure control

Kyle Alves, Senior Lecturer of Operations and Information Systems, University of the West of England.

Felix Ritchie, Professor of Applied Economics, University of the West of England.

Corresponding author: Felix Ritchie, Bristol Business School, University of the West of England, Coldharbour Lane, Bristol. BS16 1QY Felix.ritchie@uwe.ac.uk

Abstract:

Statistical agencies and other government bodies increasingly use secure remote research facilities to provide access to sensitive data for research and analysis by internal staff and third parties. Such facilities depend on human intervention to ensure that the research outputs do not breach statistical disclosure control (SDC) rules.

Output SDC can be principles-based, rules-based, or something in between. Principles-based is often seen as the gold standard statistically, as it improves both confidentiality protection and utility of outputs. However, some agencies are concerned that the operational requirements are too onerous for practical implementation, despite these statistical advantages.

This paper argues that the choice of output checking procedure should be seen through an operational lens, rather than a statistical one. We take a popular conceptualisation of customer demand from the operations management literature and apply it to the problem of output checking.

We demonstrate that principles-based output SDC addresses user and agency requirements more effectively than other approaches, and in a way which encourages user buy-in to the process. We also demonstrate how the principles-based approach aligns better with the statistical and staffing needs of the agency.

Keywords:

Confidentiality, statistical disclosure control, operations management, output checking

1. Introduction

In the last two decades, one of the key growth areas in official statistics has been the availability of confidential data for research user by academics, private sector analysts and government departments. On the demand side, users want increasing granularity in the data to address more specific policy issues. On the supply side, government data holders are under pressure to leverage their investment in data collection by maximising data use across a range of stakeholders.

Much of this data is confidential and personal, such as health or tax data. Traditionally, the privacy of respondents was managed by reducing the detail in the data, either to a level at which the data could be distributed without restriction (public use files, or PUFs), or with more detail left in the data but access limited to licensed users (scientific use files, or SUFs).

As data use has grown, so have concerns about whether the confidentiality protection is adequate. The new risks include [1] the re-identification possibilities of social media, the third-party holding of confidential data implied by the growth in administrative data as a source, and massive computing power with the ability to re-identify source data through brute force methods. Anonymization methods which were widespread and deemed adequate some years ago (for example, simple rounding of totals or cell counts of three) no longer meet acceptable standards.

Observed practice suggests five solutions: (1) increased reduction in detail; this risks making the data valueless (2) for SUFs, tighter contracts or licences; this assumes that there is a linear relationship between strict licensing conditions and user behaviour for which there is no strong evidence (3) replacing genuine data with synthetic data; this does not eliminate risk, and users are often uncomfortable about basing analysis on imputed data (4) ‘query servers’ allowing queries on the data without seeing it, and (5) research data centres (RDCs) which

allow full access to ‘secure use files’ (SecUFs) in an environment physically and/or digitally overseen by the NSI.

Table query servers, producing simple cross-tabulations and counts, are becoming widespread and effective at meeting many users’ needs for dynamic tabulations. More complex query servers offering a much wider range of analysis are now being developed, such as Statistics Norway’s elegant system at www.microdata.no. Query servers apply confidentiality checks automatically to outputs.

However, for detailed analysis researchers need access to the full microdata via an RDC. The great success story of this century for official statistics has been the use of virtual RDCs (vRDCs), where thin client technology has allowed data holders to provide the security of a physically restricted environment whilst allowing users to access the environment from more convenient locations. Many high-income countries have at least one facility operated by the National Statistics Institute (NSI) or a data archive; in the UK alone there are six general-purpose vRDCs offering the microdata underlying official statistics to a variety of users in government and academia.

RDCs manage confidentiality at the point of access but create a new risk of confidentiality breach through publication [2]. Both SUFs and SecUFs have some identification risk, so it is possible that a published output might reveal some confidential information. This risk is higher for SecUFs as the data is much more detailed. All RDCs therefore operate a system of output-checking before publication (output statistical disclosure control, or OSDC) to manage this risk.

There are two approaches to managing output-checking for conformance to regulation [3]: ‘rules-based’ and ‘principles-based’. The former sets strict rules for releasing output and applies simple yes/no criteria; the latter uses flexible rules-of-thumb and creates an environment for negotiation between researcher and output-checker. Because rules-based is very limiting in research environments, our experience is that most organisations claiming to

be rules-based operate a 'rules-based but sometimes…' system allowing for ad hoc relaxation of rules. We refer to this as 'ad hoc' output checking.

Data holders choosing between OSDC regimes consider two issues: risk, and operational efficiency. Superficially, risk should be irrelevant; all regimes operate the same statistical models for determining whether a specific output poses any risk. In practice, as is demonstrated below, rules-based regimes have a higher statistical risk but are often perceived to have lower risk.

However, our experience is that data holders are usually more focused on operational efficiency: which approach uses resources most effectively? A rules-based system can, in theory, be automated or run by humans with little statistical training. The principles-based solution requires input by humans able to discuss technical matters with researchers. Prima facie, the need for flexibility in principles-based models implies a more costly and laborious solution; this has long been established in the management literature on bespoke production [4]. In reality, the principles-based solution was designed *specifically* to reduce resource cost while also reducing risk [5], and the available evidence tends to support this.

There are two reasons for the misperception of the principles-based model as high-cost. First, data holders are often unfamiliar with the activities of research users of data, and so view them through the lens of their own outputs; these are typically tabulations which have strict rules applied for comparability across time and alternative breakdowns. Second, data holders' experience of OSDC is usually limited to the statistical literature, which focuses on arbitrary 'intruders' applying mechanical procedures to breach confidentiality (see e.g. [6]). Together, these factors encourage an over-simplistic view of the research environment which drives data-holders' perceptions of risk and benefits.

The extensive management literature on process efficiency is as robust and canonical as the statistical literature on SDC, but has played little role to date in OSDC thinking. This paper

seeks to change this by introducing a model familiar to operations management literature: that of 'runners-repeaters-strangers-aliens' (RRSA) ([7, 8]). This model can be used to differentiate inputs of demand from customers (in this case, the requests from researchers for data cleared for publication) and uses the different characteristics of those inputs to develop optimal operational responses to increase organisational readiness. Using this framework, we contrast how the rules-based and principles-based approaches address the different challenges posed by real research environments. It is then straightforward to demonstrate how the 'one-size-fits-all' rules-based model achieves neither operational efficiency nor effective risk reduction in the RDC environment. Similarly, we can also analyse why the 'rules-based-but…' ad hoc approach fails to achieve the operational advantages of the full principles-based approach.

The next section summarises the literature on the topic on output regimes and on process management. In section three we develop the output-checking problem, and in section four we show how the RRSA model can be applied to this procedure. Section five discusses empirical cost assessments. Section six concludes.

While acknowledging that many government departments produce data for re-use by researchers in academia and government, for clarity in this article we assume that the data has been collected and made available by a national statistical institute (NSI).

2. Literature review

Output checking

Output statistical disclosure control (OSDC) is a relatively new field. Until recently, the SDC literature focused almost exclusively on two problems: anonymization of microdata, and protection of tabular outputs; see for example [9], or the *Privacy in Statistical Databases* biennial conference publication. Since the development of RDCs in the early 2000s, a small

number of papers began to appear considering particular outputs such as regressions ([10, 11, 12, 13, 14], for example) as well as general guidelines for users of RDCs [14].

The concept of generalised output SDC, particularly in research environments, was proposed in 2007 [15]. Ritchie [16] introduced the concept of 'safe outputs', usually referred to now as 'safe statistics' [17] or 'high/low review statistics' [5].

Brandt et al. [18, revised edition 19] used these and operational practices to produce the first widely-available guide to OSDC for managers of research facilities. The guide was included as a chapter in Hundepool et al.'s broadly successful attempt [6] to provide an overview of state-of-the-art SDC techniques. The guide is the main source for most subsequent practitioner manuals [e.g. 20, 21, 22, 23]).

Part of the reason for the guide's popularity is its neutrality on the clearance process. [18] contains the first practitioner guide to both principles-based OSDC (PBOSDC), rules-based OSDC (RBOSDC), and the practical differences in implementation. It offered guidelines for NSIs adopting either system without demanding that either be adopted. Implicit throughout the guide, however, is that the determination of statistical risk in any particular output is independent of the clearance process: PBOSDC and RBOSDC are simply alternative ways of addressing the management of resources and risk.

A non-systematic poll of 12 RDCs [24] found that RDCs were 50-50 split between RBOSDC and PBOSDC. Discussions of the merits of the two are largely confined to practitioner meetings or papers; for example, [2] discusses how principles-based operates in an academic research network. The only peer-reviewed journal article to directly address the choice of OSDC regimes is [3]. This argues that PBOSDC is statistically and operationally superior, but acknowledge its institutional drawbacks: the principles-based model requires a greater institutional commitment to user training and active risk management, whereas the familiarity of RBOSDC makes it an easier 'sell' to data holders.

In 2017 the UK Office for National Statistics (ONS) revised the national training for UK-based researchers working with confidential microdata [5]. The ONS operates principles-based systems. The previous training model, which dominated UK training from 2004 and strongly influenced other countries' confidentiality training, treated PBOSDC as a statistical problem. The revised model was the first document to be explicit about the operational justification.

Models of user segmentation

PBOSDC implicitly acknowledges that research and researchers have multiple skills, interests and demands [15]. This becomes manageable when considering how demand inputs can be segmented. The notion that variety is introduced by different types of customer input requests, and the optimal organisational response to customer variety requires different approaches to operational delivery, is well-established in the discipline of management.

The foundations of this approach can be identified in research on improving organisational readiness through operational efficiency. Whilst exploring methods of increasing effectiveness of Just-in-Time (JIT) manufacturing strategy, Pareto analysis was applied to manufactured products to describe the demand pattern of products originally identified as "regular runners, irregular runners, and strangers" [7:486]. The categorisation was used to better understand the predictability of the customer request and its impact on having the required organisational resources in place to fulfil the order. Parnaby [7] proposed that efficiency gained through JIT success relied on a dependable stream of resources for 'runners' and 'irregular runners' (later called 'repeaters'). 'Strangers' require increased levels of customised work, making it less amenable to JIT workflow management and therefore less efficient.

While [7] does not define these labels, the terms are described in a seminal Business Process Management (BPM) paper [25]:

- Runners – demand which is part of the regular routine, predictable resource requirement

- Repeaters – intermittent and uncertain demand, some known resource requirement
- Strangers – much less predictable demand, very limited insight for resource allocation

'Aliens' were a later addition [8], describing requests from the customer which are so infrequent or unfamiliar that pre-existing knowledge is generally not applicable. Thus, a state of operational readiness for such a request cannot be achieved.

To better illustrate these terms, we provide examples drawn from the context of bank service offerings. *Runners* are requests to check an account balance or to perform a simple transfer of funds. These are standardised processes that are performed in high volumes every day. *Repeaters* might be a request to open an account in person. These requests are presented less frequently and typically require personalised service and physical document checks. The processes for these are well-documented and trained resources are readily available, but the pattern of demand is less predictable. A s*tranger* might be a rare case where a customer attempts to open an account with an unknown or unrecognised official document. In this case, because there may not be appropriate resource readily available to make an official judgement on a previously unseen form of evidence, staff may have to request specialist input. This request consumes additional resource time to simply analyse the nature of the request before a decision can be made to accept or reject. That resource may be known to the organisation and trained to make such a decision, but the readiness of the resource is negligible because the demand could not be predicted. An *alien* might be an input from the customer which challenges known capability begging the question "*Is this even possible?*" This input is unpredictable, such as if a customer requests to open an account in the name of their household pet. Such a product-offering may not exist, nor might it even be possible given the existing processes in the delivery system. This may result in an idea for a new product, but the delivery system is not currently designed to meet that customer-introduced variety.

[25] draws attention to the connection between the resource consumed in the production process and the variety in customer demand. In his view, demand variety has multiple dimensions: changes in volume and differences in requested output. This connection draws heavily on a concept especially relevant here, Ashby's 'Law of Requisite Variety' [26]. Requisite variety mandates that any system must meet request variety with a similar variety in production capability; or else it must attenuate/reject that request to remain viable. Thus, the success or failure of a delivery system is determined by its adequacy in aligning operational capability with the network of customers and suppliers [27, 28].

The categorisations of demand characteristics act as an aid to the organisation in managing its environment and maintaining viability through the efficient allocation of resource. In this way, efficiency is the product of how well the delivery process is designed to meet the variety in demand.

Alignment between the design and the context in which it will operate has been shown to lead to optimal performance [29, 30]. Similarly, research has identified a connection between design and performance, whereby "inadequate service design will cause continuous problems with service delivery" [31]. Considering the potential applications in the context of the ONS, the research [32] is highly relevant: in the face of higher request variety, an organisation can employ a design strategy which uses different operational means of delivering similar outputs to customers.

This concept was empirically explored in [33] where complexity of customer demand was shown to determine the level of customisation provided by the delivery system. This approach uses parallel processes that are each designed to deliver the same output but employ different levels of customisation. Highly standardised, efficiency-focused processes are put in place for 'runners', while other processes use increasingly higher levels of customisation to enable the organisation to react to the 'strangers'; increasingly complex customer inputs. The unfamiliar

nature of ‘aliens’ may require innovation in process design in order to accept the related presented variety.

Encountering ‘strangers’ and ‘aliens’ forces an organisational choice of whether to accept the input request or attenuate the variety and reject the request. If accepted and produced, the new output may then be offered to other customers by continued implementation of the newly created process [8].

Conversely, the organisation may implement design which requires greater participation by the customer in the creation of the output. Frei [29] suggests the accommodation of customer-presented complexity through ‘low-cost accommodation’. By shifting work away from the organisation and back to the customer, the organisation can derive some benefit from efficiencies in resource allocation. In this case, customers are given access to the delivery system in order to ‘self-serve’ and create their own outcomes. The reframing of the customer as a co-producer of value was researched in the marketing discipline [34], and later from an operational perspective [35] in the concept of value co-creation.

There may be questions about whether users will adjust their demands in order to create conformity with the parameters which would make their request a ‘runner’ or not. The creation of these co-production processes creates mutually positive goals for both the customer and the organisation [36]. This incentivises the user to participate in the co-production activity. By fully understanding and adhering to known parameters, they can generate their specific, unique outcome with minimal resource-consuming friction with the organisation.

Sufficient evidence exists to support the application of the RRSA model to OSDC for the purposes of increasing efficiency in the use of resources through adjustments to the organisational delivery system. Central to this, it is necessary to explore the alignment

between the nature of the request from the customer and the characteristics of the process required to fulfil that request.

3. Rules-based, principles-based and ad-hoc output-checking

Figure 1 below shows a typical output-checking process from a secure environment managed by a National Statistical Institute (NSI):

[Figure 1 here]

The researcher works in an environment where he or she cannot directly take away statistical results (note: some facilities allow more 'trusted' users to check and release their own outputs). The researcher places the outputs to be released in some predefined location in the secure environment and asks the support team to check and release the output. The support team can extract results from the secure environment. If the support team decides the output is non-disclosive, it sends the results out to the researcher's (open) home environment.

For expository purposes, we will assume that the researcher has asked for a frequency table to be released, and that the support team operates a simple threshold rule of three; that is, the table must have at least three observations underlying each cell in the table. So, Table 1 below, part (a) passes the SDC rule but part (b) does not:

[Table 1 here]

Under a rules-based approach (RBOSDC), this is a hard limit; no exceptions are allowed. Setting the rule is problematic as the rule is trying to balance both confidentiality and utility. Too high a value prevents the publication of useful but non-disclosive findings; too low a threshold allows many useful results to be published but increases the risk that disclosive results leak out.

Under the principles-based approach (PBOSDC), the rule becomes a 'rule-of-thumb': it guides decisions but is not always followed, and can be adjusted up and down as necessary. The researcher can argue [5] that the rule can be ignored if, and only if:

- the output is non-disclosive, and
- the detail in the output is important to the researcher, and
- this request for an exception is a rare occurrence for the researcher

The first condition is obvious. The second condition ensures that the output-checker and researcher only spend time negotiating over an output when the result matters to the researcher. This is appealing to researchers as it puts them in charge of deciding when something is 'important', rather than the output checker. Thus, Table (b) above could be released if the researcher demonstrated that the small value was non-disclosive and essential for publication. The third condition ensures that researchers do not abuse the system. Note that the terms "important" and "rare" are not specified – this is an area for the researcher and output-checker to negotiate [5]. As a result, training the researcher to understand the concept is necessary; effectively, this is the model of low-cost accommodation [29].

PBOSDC is two-way: the output-checker can also argue that the rule-of-thumb is inappropriate in a specific case because it does not protect confidentiality. For example, in the above case the output-checker may argue that a higher threshold is needed because the data are particularly sensitive and the patients are easily identified. Some organisations (for example, National Records for Scotland) operate a two-tier system with a lower 'regular' threshold and a higher threshold for outputs based on more sensitive data.

PBOSDC systems usually use higher thresholds (10 is common) than RBOSDC. Use of an overly restrictive rule-of-thumb does not limit research as the researchers always have the opportunity to argue for an exception. Hence, the threshold can be set high as it only has to address the confidentiality problem; the utility problem is dealt with by the exception

mechanism. The rule-of-thumb threshold is set at a level which provides a high degree of confidence in the non-disclosiveness of results while simultaneously not leading to many requests for exceptions. In practice, genuine research environments can tolerate much higher thresholds under principles-based models.

It is the combination of stricter rules-of-thumb and the ability to use discretion in applying those stringent rules that gives the principles-based approach its superior risk management. Under RBOSDC, a single rule must do two jobs: protect confidentiality (by having a higher threshold, for example), and allow useful, non-disclosive output to be published (which is limited by having a high threshold). Security and efficiency must be traded off. In contrast, under PBOSDC, the rule has one job (protect confidentiality in most cases); efficiency comes through negotiation when it matters.

Rules-based models also fail to provide the imagined guarantees over security. Consider the following tables:

[Table 2 here]

In (c), all cells have at least 5 underlying observations. However, it is clear that an *implicit* table is being generated: the complement to the proportion with the genetic marker is the proportion without it. Part (c) shows that there are 2 males (10% of the 20 in total), diagnosed with diabetes in the dataset who have the genetic marker.

Part (d) shows the problem of *class disclosure*. All males in this dataset diagnosed with diabetes have a BMI greater than 25 i.e. they are overweight or obese. It doesn't matter that there are twenty individuals in this group, well above the threshold; something is now known about *all* males with diabetes diagnosed.

These pedagogical examples illustrate some limitations in a simple threshold rule; other examples can be developed. It can be argued that the rules-based organisation simply needs to

develop rules to cover these cases. However, as [15] demonstrates, developing complex rules to cover special cases quickly becomes unwieldy compared to the principles-based approach of simple rules-of-thumb and flexibility in interpretation. When combining the higher thresholds used in PBOSDC with the need to explicitly define what happens in many states of the world in a rules-based model, it is clear that RBOSDC is the higher-risk option.

If this is the case, why do risk-averse organisations use RBOSDC? Two reasons are invariably given.

First, rules are said to be simpler for everyone to use (researchers and output-checkers) and easier to explain to data-holders who want to be reassured when depositing their data. The latter is a valid point: data holders, if aware of SDC at all, are likely to be familiar only with the traditional model of SDC for tabular data in a hostile environment. Simple rules reflecting that knowledge have immediate appeal, even though the sense of security in the familiar is not warranted.

The second, and more common, reason given is that RBOSDC uses fewer resources: applying simple rules should be easier and require lower-skilled operators than a system which leaves open the possibility of negotiation over any statistical artefact. Principles-based systems cannot be less resource-intensive than rules-based models in the absence of queries and must require more resources if the checking staff must deal with queries. Moreover, those resources involve output-checkers with statistical skills, which are not necessary for the rules-based system.

There is an alternative to RBOSDC and PBOSDC which is widely implemented. Almost no rules-based organisations operate in the simplistic way described above. All have some informal arrangement allowing researchers to argue that outputs which break the rules can be released in certain circumstances. We will refer to this as ‘ad hoc’ output SDC (AHOSDC). This method

can provide some of the flexibility/efficiency gains of the full principles-based approach without the potential free-for-all.

At first glance, this approach seems to offer the best of both worlds. In practice, it suffers the key problems of both. First, it does not address the risky nature of rules-based by making no allowance for output checkers ignoring rules to block disclosive outputs. More importantly, not formally acknowledging that rules are flexible can create a lack of clarity, causing uncertainty and inefficiency. It also makes resource allocation harder: should output checkers have statistical skills when the formal policy of the organisation says that they do not need them?

Some organisations argue that the simple threshold rule presented above is a straw man: more complicated rules can achieve both the security and the flexibility of PBOSDC. Unfortunately, this is extremely difficult to do in genuine research environments. Ritchie (2007) provides a counter example where a simple, specific, unambiguous, 17-word threshold rule rapidly becomes a woolly 47-word mouthful which requires specialist interpretation, and which is easily challenged by a researcher wanting to make a point.

Fundamentally, the reason why RBOSDC (and AHOSDC) fails to meet the twin targets of efficiency and security is because it is grounded in the SDC literature which sees this as a statistical problem generated by an arbitrary 'user' type of individual. In contrast, PBOSDC sees output checking as a process problem, caused by multiple types of client and client needs. To see why this makes such a difference in implementation, we now turn to the management literature.

4. Output checking as a user segmentation problem

The process perspective on output checking starts from the recognition that different types of researchers, and types of output, produce different demands on the NSI. As noted in the literature review, the concept of 'requisite variety' was established as far back as 1956, and

there is a well-established management literature which uses segmentation of customer demand variety as a way of efficiently allocating resources to better achieve operational readiness. Simpler demands are automated as far as possible, leaving specialist resources to be concentrated on the more specialist and typically more complex high-value cases.

We employ the 'runners, repeaters strangers and aliens' (RRSA) model. Using the RRSA terminology, we can divide output requests into four separate types, with some indication of how often these occur (based on personal experience):

[Table 3 here]

Like the example made earlier of checking account balances in a bank, the runners are the bread-and-butter of microdata research. They include simple descriptive statistics such as mean of the observations, or frequency counts in categories, which are usually presented with high numbers of observations. The runners also include 'safe' (or 'low review') statistics, such as regression coefficients, where there is no meaningful disclosure risk [36]. In theory, these outputs could be reviewed automatically using the programs tau-Argus or sdcMicro, for example. In practice, they are manually reviewed as this is faster.

The repeaters are the outputs which require the reviewer to make a judgement based on context. This aligns to the example provided earlier described the case where bank staff evaluates the official documentation for opening an account. At the NSI, a scatter plot of regression residuals might be submitted; the checker would want to evaluate the risk in any outliers. Alternatively, this could be a simple table with counts below the threshold (as in a PBOSDC 'exception' request); the checker is then being asked to make a judgment on whether this is non-disclosive, infrequent, and important.

These two cover almost all outputs from research centres. Note that in RBOSDC, only the runners exist: an output cannot be cleared unless a known unambiguous rule exists. For

PBOSDC, allowing for repeaters is essential: this flexibility to review outputs in context allows much more restrictive (that is, protective) rules to be placed on runners.

The strangers are where the researcher produces something that the output checking team hasn't seen before. This aligns to the earlier example where a bank customer produces unrecognised but official documentation to support opening an account. It could be a novel output (for example, being asked to make a decision on a Herfindahl index for the first time), or familiar outputs presented in an unfamiliar way (in one case, a project which required a very large number of intersecting tables). These require multiple skills: a reasonable degree of statistical knowledge, an ability to judge evidence effectively, and the social skills to hold productive discussions with the researcher.

This highly skilled resource is expensive; analytical work is the most efficient utilisation of that resource, and so limiting the time spent by that resource on checking outputs is important for the organisation. Ideally, a stranger only appears once: the role of the reviewer is to decide whether this specific output is to be released, and how future outputs of the same type should be classified. For example, on first encountering a heat map or box-and-whisker plot, the former would be classified as a runner, the latter as a repeater. In the bank analogy, once the unrecognised documentation has been reviewed, this would be expected to generate a rule as to whether similar documents are valid in future.

Finally, the aliens are those outputs for which the facility was not designed; for example, the release of a linked dataset constructed by a researcher. This does not require statistical knowledge at all, but rather understanding of the purpose of the facility. It may lead to a redesign of the facility (say, an isolated section to allow linking to take place).

An amended model of the process is presented in Figure 2, which reflects the various process flows associated with a triage activity sorting runners, repeaters, and strangers.

[Figure 2 here]

As well as managing resources, this structure also allows the NSI to make the researcher an active part of the clearance process; a co-production approach which utilises 'low cost accommodation' [29]. PBOSDC training for researchers [5] emphasises three points:

- Runners are done quicker than repeaters
- If the planned output is a 'stranger', involve the clearance team as soon as possible to comment on the likely acceptability and avoid time wasted on unreleasable outputs
- Provide all the information for the reviewer to put the output into one of the classes

The aim is to make the researcher see that his or her behaviour has a direct impact on clearance times, and to provide the researcher with incentives to improve them. By making the researcher a co-producer in the clearance chain in carrying out the preparatory work, the output-checker finds his or her workload reduced. The awareness of researchers that they can directly affect response times builds a feeling of control and hence engagement in the process (see [38] for a detailed review of community-based training, and [5] for delivery). Consider Table 4 below.

[Table 4 about here]

Table 4 reveals the impact the RRSA classification has on the need for staff resources to review outputs. Of the eight examples provided, five of the output types classified as runners can be managed with little-to-no staff resource consumption. This has the obvious efficiency impacts for the organisation, while also providing the user a simpler, faster ability to generate the data reports they require. The approach of co-production permits the user to shape their input in such a way as to make them 'runners' produced with little organisational friction. However, to fully exploit this, the user also needs to be aware of their role, and hence this sort of discussion is central to PBSDOC training such as [5].

The RRSA model also provides a clear structure for staff resources. Consider the skills needed by an output-checker for the different types of output, and that person's discretion to ignore the rules of thumb:

[Table 5 here]

This creates a hierarchy of technical skills allowing different staff to be allocated to different roles. It also simplifies skills acquisition and staff training, by providing a clear path to personal development based upon experience and knowledge of the data.

This differentiation of skills is the second reason why apparently lower-cost models of output clearance fail to achieve the operational gains of PBOSDC. For RBOSDC, only runners and strangers should exist: there are fixed rules, which may be added to as new statistical products occur. This implies that the bulk of the work can be carried out by checkers with minimal training in statistics or data.

However, the systems run by most NSIs are ad hoc; that is, notionally the rules are hard but in practice researchers ask for, and get, some flexibility. The flexibility may depend on the data, the statistic, or sometimes whether the researcher is 'trusted' or not. The flexibility is important: without it, the NSI is likely to lose the goodwill of the researcher. The difficulty is that, because the flexibility is not officially sanctioned, it is less clear whether a clearance is going to be simple or complex, allowed or blocked. Clearance times become less certain; and because any clearance might be an exception, all clearance staff need to have the ability to handle exceptions. The efficiency gains from having clearly delineated production processes have been lost.

5. Implementation issues

Cost-effectiveness and resource use

PBOSDC was devised and first implemented at the UK Office for National Statistics (ONS). While the statistical benefits became evident over time, the initial appeal was as a way of keeping overall costs down via low-cost accommodation. In this, it appeared to be successful. At its initial peak in 2008-2010, the ONS' secure research facility was reportedly releasing more outputs from more researchers at significantly lower staff cost than comparable European facilities[1].

However, there are extremely efficient rules-based systems. Statistics Norway runs both a full-service RDC, and a remote job model (microdata.no) developed in collaboration with the Norwegian Center for Research Data (NSD)[2]. Microdata.no takes the rules-based model to its logical conclusion: all decisions are computerized, and the system does not allow outputs which do not have a clearance rule attached. This automated approach is resource-efficient. Importantly, it allows users to see release decisions in real time and so, despite the very strict rules (for example, the initial threshold for tables is set at minimum of 1,000 observations), user responses have been very positive. The model has attracted substantial interest from NSIs and other organisations.

It is also feasible to operate very cost-effective ad hoc systems. In social science, the pre-eminent example is LISSY[3], which has been running for almost two decades and allows users to submit code to run analyses of the Luxembourg Income/Wealth Studies. Like microdata.no, simple strict rules are applied automatically by the server, but computerized triaging allows the option 'set for review' (that is, send to a human for checking). As a result, the rules can be

[1] This information was gained in conversations at the Eurostat expert group 2009-10, and from presentations by RDC operators at this time.
[2] https://microdata.no/
[3] https://www.lisdatacenter.org/data-access/lissy/

less stringent than microdata.no. As in Norway, users get immediate feedback on whether code is allowed to run or not, and users are encouraged to recode rather than waiting for review. Despite the small staff, LISSY handled 73,000 job requests in 2018 and has shown continual growth in both user numbers and data requests, indicating a high level of user satisfaction.

Thus, there are examples of efficient principles-based, rules-based, or ad hoc OSDC systems. Applying the RRSA model to the different examples helps us understand this efficiency. The Norwegian model only has runners and aliens; the latter are used to identify new rules to widen the class of runners. With only runners, automatic clearance of all outputs is the logical and cost-effective conclusion. In the case of LISSY, repeaters are allowed but strongly discouraged via immediate feedback. Just as for PBOSDC, this feedback is designed to encourage users to change their behaviour.

Organizational limits

Methodological problems limit the chance of comparative evaluation, but it is clear that no solid evidence exists to support the argument that PBOSDC is more expensive than other solutions. On the contrary, seeing RDC disclosure risk management from an operational perspective makes clear that PBOSDC is expected to be more efficient than both ad hoc solutions and rules based solutions, albeit for different reasons for the two approaches.

Although a compelling theoretical case can be made for PBOSDC in research environments, two institutional factors can limit its adoption and its effectiveness.

First, cost streams differ. PBOSDC requires initial investment in training both users and output checkers to fully achieve the efficiency gains [38]. A completely automated rules-based system like microdata.no also requires a great deal of planning. In contrast, an ad hoc approach can be implemented with little user training, simple rules, and output checkers training on-the-job. For a low-volume clearance system, the ad hoc model may be the most cost-effective solution.

The expected growth in demand also affects the cost stream: growing systems allow initial investments to be leveraged more effectively.

Second, changing institutional practice can consume time and resources. Older clearance systems are, in practice, likely to be ad hoc. Familiarity creates organisational inertia which may be impossible to overcome. Our own experience is that inertia is often a key stumbling block to introducing new processes; and the more complex or specialist the area of change, the more familiarity becomes a reason for resisting change.

6. Conclusion

Secure research access to the most sensitive microdata has been one of the great success stories for NSIs this century. It came from realising that simply reducing data detail was a dead-end. Instead, novel ways of working with researchers opened a range of options. For all of these new ways of working, output-checking is a key part of the operational system.

Perceptions of output-checking have been dominated by the statistical literature, which is designed to address the safe production of statistical aggregates. The rules-based approach is well-suited to statistical aggregates, but not research outputs. Hence OSDC was born as a field, with PBOSDC its standard-bearer. But to those brought up on the traditional statistics, PBOSDC seemed an operational nightmare: how can an explicitly 'flexible' (i.e. uncertain) world, requiring greater statistical understanding and more training for everyone, be both safe and scalable?

When viewed from an operations management perspective, the answer is that the efficiency gains come precisely from the elements that worry traditionalists. Tough, simple rules with flexibility at the margin for high-value outputs means that the 80%-90% of runners can be handled quickly by automated processes or staff with minimal training. Allowing researchers to choose when their runners become repeaters saves the output-checker carrying out this

function. This is an example of both Frei's low-cost accommodation [29] and customer co-production [34]: the NSI has effectively turned the customer into part of the workforce. More importantly, it gives the researcher some control over the process, and so builds engagement. There are upfront training costs, but these should be seen as investment expenditure.

From the management literature, there are no surprises that a one-size-fits-all model allocates resources less efficiently than a segmented-markets model; nor that the latter is better at exploiting customer self-service. This 'requisite variety' has been a core of management thinking for over half a century. This operational lens can also shed light on how some very efficient rules-based or ad hoc systems manage their resources so well: substantial upfront investments are supported by instant feedback mechanisms. These ensure that expectations are managed effectively so that customers can self-select into the 'runners' category.

What is perhaps less obvious is that this also produces better statistical outcomes: PBOSDC is inherently lower-risk than RBOSDC. In theory, the disclosure risk in an output is independent of the checking process: the criteria for assessing disclosiveness are statistical, not operational. This is only true if resources for checking are unlimited. In practice, the checking process does affect statistical outcomes.

First, rules-based systems have to trade off confidentiality protection and utility in a single rule, but in PBOSDC the rules-of-thumb are there just to manage confidentiality; utility is managed separately by the exception mechanism. Ceteris paribus, a PBOSDC system operates stricter statistical criteria then its rules-based equivalent while providing greater utility.

Second, PBOSDC concentrates output checker resources on important and rare cases, rather than spreading those resources across all outputs. As a side-effect, PBOSDC also reduces dissatisfaction amongst users, a known risk factor for restricted-access systems [39].

This illustrates a wider issue. The traditional focus on statistical measures of risk, without considering the implications of operational choices, has been strongly criticised (e.g.[39]) as risky and inefficient. [40] also suggested that this is doomed to failure in a big data/machine learning world. In contrast, research into operations has much to say about effective risk management, particularly in relation to digital services (such as the 'data supply chain' model [41]. A change in emphasis from statistical to operational models of risk drawing on the extensive management literature (as in ONS' 2019 course for output checkers which uses the SSRA framing) should help NSIs to improve delivery on the joint objectives of security, efficiency, and customer service.

References

[1] Statistics Authority (2018) National Statistician's Quality Review on Privacy and Confidentiality. Ed. G. Roarson. UK Statistics Authority, December.

[2] Lowthian P. and Ritchie F. (2017) *Ensuring the confidentiality of statistical outputs from the ADRN*. Technical report no3. Administrative Data Research Network

[3] Ritchie F. and Elliot M. (2015). Principles- versus rules-based output statistical disclosure control in remote access environments. IASSIST Quarterly 39:5-13

[4] Chase, R. (1981) "The Customer Contact Approach to Services…" Operations Research, 29, 4.

[5] Office for National Statistics (2019) Safe Researcher Training [2017 onwards]. Office for National Statistics, Research Support Service. Last viewed June 2019.

[6] Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Lenz, R., Longhurst, J., Schulte Nord-holt, E., Seri, G. and De Wolf, P-P. (2010). Handbook on Statistical Disclosure Control. ESSNet SDC. http://neon.vb.cbs.nl/casc/.\SDC_Handbook.pdf

[7] Parnaby, J. (1988). A systems approach to the implementation of JIT methodologies in Lucas Industries. The International Journal Of Production Research, 26(3), 483-492.

[8] Aitken, J., Childerhouse, P. and Towill, D., 2003. The impact of product life cycle on supply chain strategy. International Journal of Production Economics, 85(2), pp.127-140.

[9] Willenborg L. de Waal T. Statistical Disclosure Control in Practice: v. 111. Spinrger: New York: Lecture Notes in Statistics; 2013

[10] Reiter J. (2003). Model diagnostics for remote-access regression servers. Statistics and Computing. 13:371–380

[11] Reznek, A. (2004) Disclosure risks in cross-section regression models, mimeo, Center for Economic Studies, US Bureau of the Census, Washington

[12] Reznek A. and Riggs T. (2005) Disclosure risks in releasing output based on regression residuals. ASA 2004 Proceedings of the Section on Government Statistics and Section on Social Statistics pp1397-1404

[13] Ritchie F. (2006) Disclosure control of analytical outputs. Mimeo: Office for National Statistics. Edited and reprinted as WISERD Data and Methods Working Paper no. 5 (2011).

[14] Corscadden, L., Enright J., Khoo J., Krsnich F., McDonald S., and Zeng I. (2006) Disclosure assessment of analytical outputs. Mimeo, Statistics New Zealand, Wellington.

[15] Ritchie F. (2007) Statistical disclosure control in a research environment. Mimeo, Office for National Statistics. Edited and reprinted as WISERD Data and Methods Working Paper no. 6 (2011).

[16] Ritchie F. (2008) Disclosure detection in research environments in practice. In: Work session on statistical data confidentiality 2007, Eurostat; pp399-406

[17] Ritchie F. (2014) Operationalising safe statistics: the case of linear regression. Working papers in Economics no. 1410, University of the West of England, Bristol. September

[18] Brandt M., Franconi L., Guerke C., Hundepool A., Lucarelli M., Mol J., Ritchie F., Seri G. and Welpton R. (2010), Guidelines for the checking of output based on microdata research, Final Report of ESSnet Sub-group on Output SDC http://neon.vb.cbs.nl/casc/ESSnet/guidelines_on_outputchecking.pdf

[19] Bond S., Brandt M., de Wolf P-P (2015) Guidelines for Output Checking. Eurostat. https://ec.europa.eu/eurostat/cros/system/files/dwb_standalone-document_output-checking-guidelines.pdf

[20] Eurostat (2016) Self-study material for the users of Eurostat microdata sets. http://ec.europa.eu/eurostat/web/microdata/overview/self-study-material-for-microdata-users

[21] Statistics NZ (2015). Microdata Output Guide (Third edition). Statistics New Zealand. Available from www.stats.govt.nz.

[22] O'Keefe C., Westcott M., Ickowicz A., O'Sullivan M. and Churches T. (2015) Guidelines for Confidentiality Protection in Public Health Research Results. CSIRO.

[23] Griffiths E., Greci C., Kotrotsios Y., Parker S., Scott J., Welpton R., Wolters A. and Woods C. (2019) *Handbook on Statistical Disclosure Control for Outputs*. Safe Data Access Professionals Working Group. https://doi.org/10.6084/m9.figshare.9958520

[24] Australian Department of Social Services (2016) Data Access Strategy: final report. Australian Department of Social Services, June.

[25] Armistead, C. (1996). Principles of business process management. Managing Service Quality: An International Journal, 6(6), 48-52.

[26] Ashby, W. R. (1956), An Introduction to Cybernetics, Chapman & Hall Ltd., London.

[27] Beer, S. (1984). The viable system model: Its provenance, development, methodology and pathology. Journal of the operational research society, 35(1), 7-25.

[28] Pickering, A. (2002). Cybernetics and the mangle: Ashby, Beer and Pask. Social Studies of Science, 32(3), 413-437.

[29] Frei, F. X. (2006), "Breaking the Trade-Off Between Efficiency and Service", Harvard Business Review, 84, 11, pp. 92-101.

[30] Sampson, S. E. and Froehle, C. M. (2006), "Foundations and Implications of a Proposed Unified Services Theory", Production and Operations Management, 15, 2, pp. 329-343.

[31] Gummesson 1994 Relationship marketing: from 4Ps to 30Rs. Stockholm University, Stockholm.

[32] Sousa, R. and Voss, C. A. (2006), "Service Quality in Multichannel Services Employing Virtual Channels", Journal of Service Research, 8, 4, pp. 356-371.

[33] Ponsignon, F., Smart, P. A., & Maull, R. S. (2011). Service delivery system design: characteristics and contingencies. International Journal of Operations & Production Management, 31(3), 324-349.

[34] Wikström, S. (1996). The customer as co-producer. European journal of marketing, 30(4), 6-19.

[35] Prahalad, C. K., & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. Journal of interactive marketing, 18(3), 5-14.

[36] Payne, A.F., Storbacka, K. & Frow, P. (2008) Managing the co-creation of value. Journal of the Academy of Marketing Science, 36, 83–96.

[37] Ritchie F. (2019) Analysing the disclosure risk of regression coefficients. Transactions on Data Privacy 12:2 117-144

[38] Green E., Ritchie F., Newman J. and Parker T. (2017) Lessons learned in training 'safe users' of confidential data. In: Worksession on Statistical Data Confidentiality 2017. Eurostat.

[39] Hafner H-P., Lenz R., Ritchie F., and Welpton R. (2015) Evidence-based, context-sensitive, user-centred, risk-managed SDC planning: designing data access solutions for scientific use. In: Worksession on Statistical Data Confidentiality 2015, Eurostat.

[40] Ritchie F. and Smith J. (2018) Confidentiality and linked data. In Statistics Authority [1]

[41] Spanaki, K., Gürgüç, Z., Adams, R., & Mulligan, C. (2018). Data supply chain (DSC): research synthesis and future directions. International Journal of Production Research, 56(13), 4447-4466.

Tables

Table 1 Pedagogical example of a threshold rule

(a) Age versus diabetic status

		Age group	
		18-24	25-29
Men	Diagnosed	11	9
	No diagnosis	349	407
Women	Diagnosed	12	14
	No diagnosis	267	299

(b) Gene marker vs diabetic status

		Genetic marker	
		Yes	No
Men	Diagnosed	18	2
	No diagnosis	72	684
Women	Diagnosed	21	5
	No diagnosis	64	502

Note: all data fictional and for illustrative purposes only

Table 2 Pedagogical examples of problematic tables

(a) Proportion with no genetic markers

		Number	No genetic Marker
M	Diag.	20	90%
	No diag.	756	10%
F	Diag.	26	81%
	No diag.	566	11%

(b) Diabetes diagnosis versus BMI

		Body mass index			
		<18	18-25	25-30	>30
M	Diag.	0	0	3	17
	No diag.	110	511	94	41
F	Diag.	3	3	4	16
	No diag.	46	449	56	15

Table 3 Dividing request examples into RRSA types

Type (% of all input requests)	Characteristics	Example	Resource need
Runners 80%-90%	Outputs that could be checked automatically	Small simple tables exceeding rules of thumb; regression coefficients; concentration indexes	Checks for classification and yes/no rules, with an assumption of clearance
Repeaters 10%-20%	Outputs that require human but non-technical review	Multiple linked tables, large or multidimensional tables; graphs; tables where the numbers fall below the threshold	Simple tests applied to provide assurance; assumption of clearance given context
Strangers 1%-2%	Outputs requiring technical review and the development of new guidelines	New statistical outputs with no current guidelines; datasets with very unusual characteristics	Detailed review by technical staff plus development of new guidelines
Aliens n/a	Outputs not normally considered as relevant to this environment	Release of record-level data rather than statistics; release of qualitative data e.g. quotes or video images	Review of appropriateness of environment

Table 4 Example classification of specific outputs

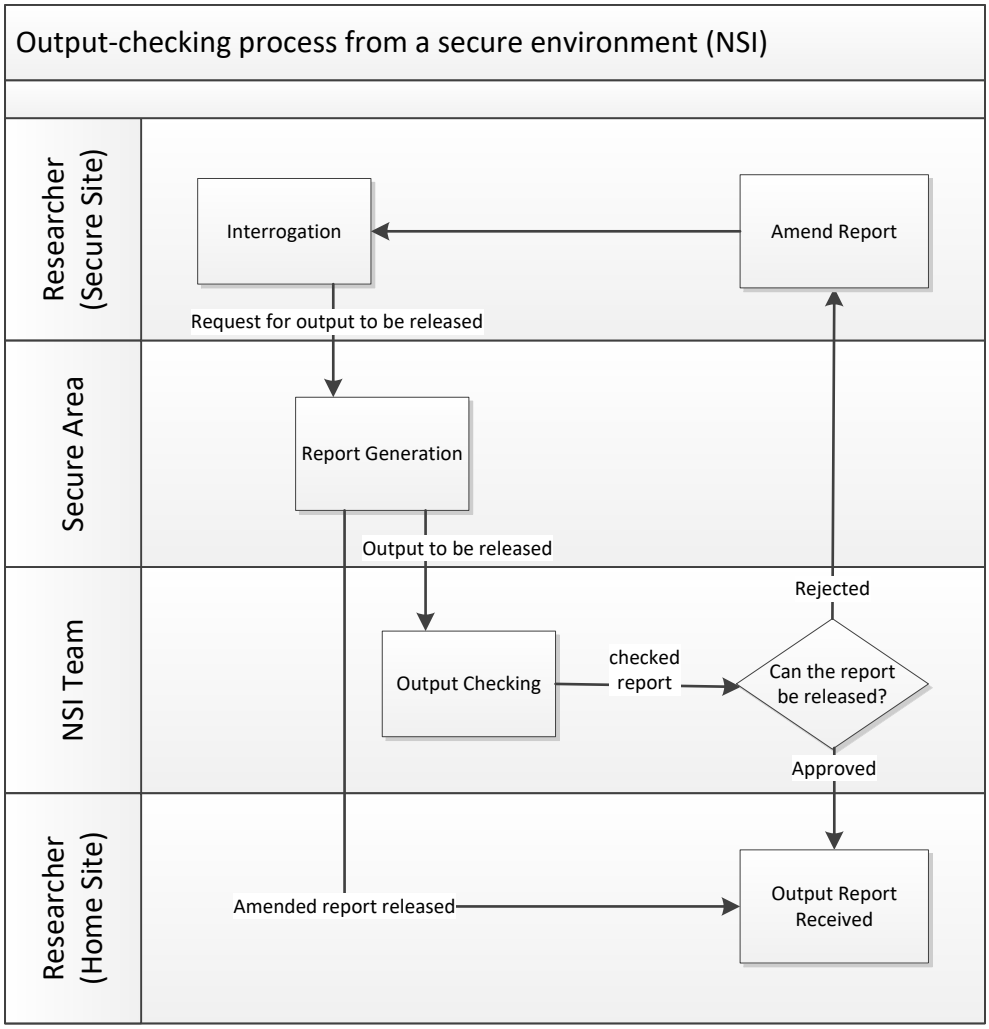
Output requested	Type of statistic	Meets criteria?	Exception requested?	Classification	Decision
Frequency table	unsafe	Yes (above threshold)	no	runner	approve
Frequency table	unsafe	No (below threshold)	no	runner	reject
Frequency table	unsafe	Yes (below threshold)	yes	repeater	review
Regression output	safe	Yes (sufficient degrees of freedom)	n/a	runner	approve
Regression output	safe	No (insufficient degrees of freedom)	n/a	runner	reject
Residual plot	unsafe	n/a (none defined)	no	runner	reject
Residual plot	unsafe	n/a (none defined)	yes	repeater	review
Tobin's Q	undefined	n/a	n/a	stranger	review and generate classification

Table 5 Output types, skill sets and discretion

Type	Skills	Rules of thumb	Standardisation
Runners	Ability to recognise types of output and follow rules	Follow	Highly standardised
Repeaters	Good understanding of data and practical (not theoretical) understanding of disclosure risk; statistically competent but not expert	Follow with interpretation	Mixed standardisation and discretion
Strangers	Statistical/data skills to understand new types of problems and take decisions	Develop new ones	High use of discretion
Aliens	Strategic perspective on operations	Out of scope	n/a

Figure captions

Figure 1 Example output-checking process



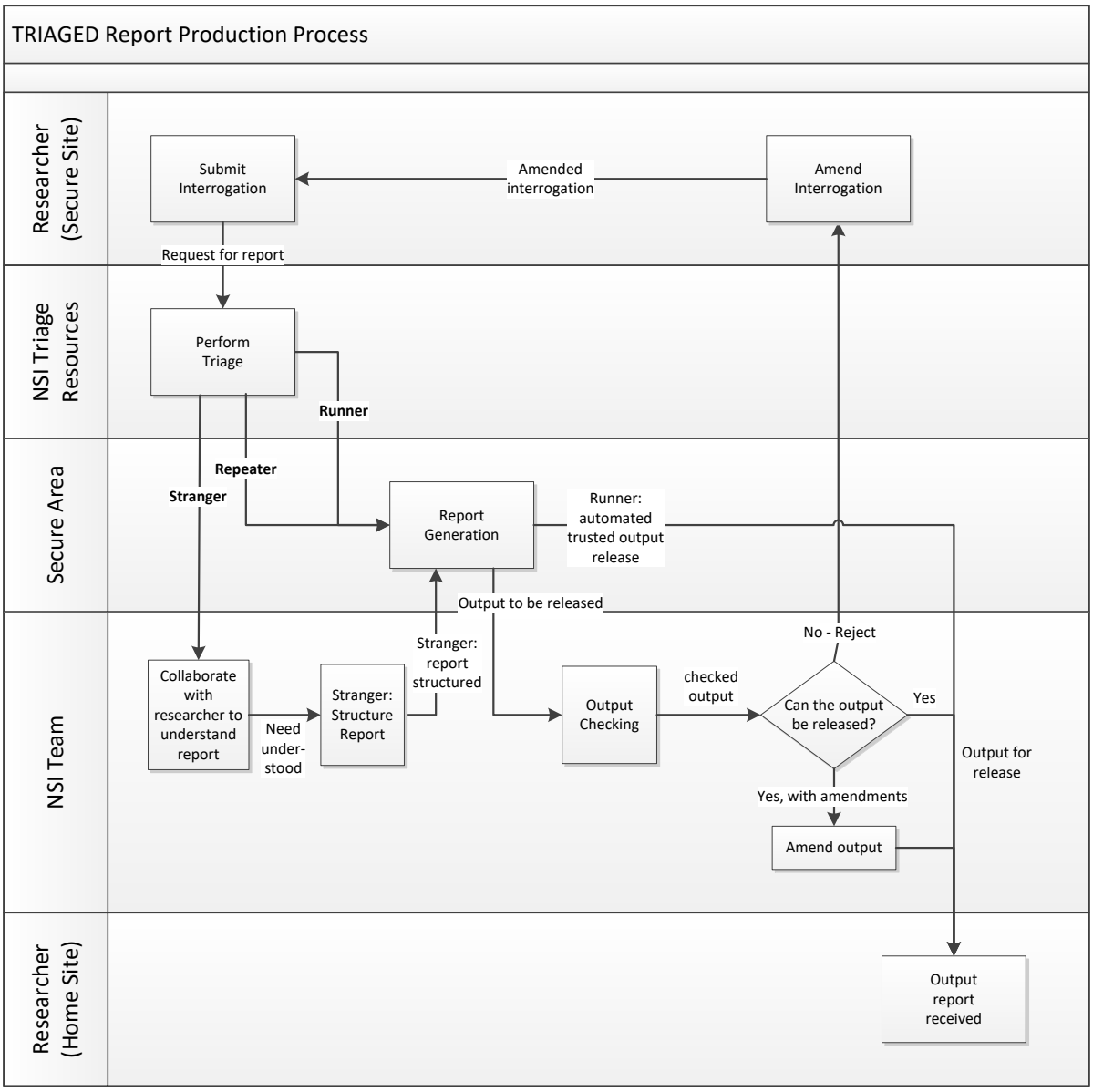


Figure 3 Output checking when viewed as a multi-stage triaged process