

1. **Executive summary**

- Algorithms to assist clinical decisions are by themselves effective at improving mental health outcomes compared to an unassisted clinician. However in practice they are rarely used.
- Designing effective algorithmic tools therefore depends both on clinical utility and addressing barriers to uptake.
- The current design of algorithmic tools is based heavily in academic research, which has limited applicability in practice and does not take issues of usability into account. This makes the majority of them unfeasible in public services because they will not be used by clinicians.
- Tool feasibility can be radically improved with proper design, a broad evidence base, and stakeholder involvement.
- Risks from algorithms were associated more strongly with inappropriate recommendations due to limited evidence, such as on individual patient preferences, organisation priorities, or contraindications for therapy.
- Risks to uptake were linked more to clinician perceptions of low utility or poor organisational fit.
- Risks could be mediated by involving clinicians in the design of the systems, and by creating systems that worked alongside clinicians when making recommendations (rather than imposing rigid decisions).

2. **Background to the submission**

2.1. This evidence has been prepared as part of a thesis submission for a Doctorate in Clinical Psychology at University College London. The full research review and protocol are available on request. The author is making the findings of their research available to the Science and Technology Committee in the hope it will be of benefit to their inquiry, particularly on the use and risk of algorithms in mental health decision-making. The author is currently completing their Doctorate with sponsorship from the NHS.

3. **Recommendations for action**

- 3.1. All algorithmic systems should be based on a stakeholder design process that incorporates relevant expert opinion, local service procedures, and organisational priorities into the decisional model, in addition to the research base. This will improve effectiveness and uptake than if grounded in the research base alone.
- 3.2. No system should be deployed without being individualised to its service context, as in 3.1. This will reduce barriers to adoption.
- 3.3. All systems should undergo a reasonable level of user testing with clinicians and service users (as appropriate) before being finally deployed. This will address ergonomic issues and other barriers to usability.
- 3.4. All systems should incorporate a process of continuous development in order to adapt it to new research, procedures, and organisational contexts as necessary. This will maintain its value over time within a given setting.

3.5. Clinician perceptions of system usefulness are not related to uptake or demonstrable value, and should not be relied upon as an indicator of either in future research.

4. Introduction

4.1. Humans have long been recognised as poorer decision makers compared to their algorithmic counterparts (for example Meehl, 1954; Clark, 1992)□, and are particularly bad at judgements under conditions of uncertainty (Tversky & Kahneman, 1973)□, such as predicting the future. While the value of algorithms is thus well established, risks from their use remain unclear, and reasons for their lack of widespread adoption is unknown.

4.2. The findings below are the result of a review of the research literature on the application of algorithms as predictive clinical decision support systems (PCDSSs) to improve mental health outcomes. A PCDSS is understood as any rule-based system involving calculations performed by a machine (or could feasibly be done so) in order to prospectively determine who can benefit from a particular intervention. This can include recommendations for therapy based on analysis of wellbeing scores, suggestions for further assessment after a new diagnosis, and other such proactive measures.

4.3. The review examined evidence for PCDSS (1) effectiveness, (2) feasibility in practice, and (3) risks. Conditions for effectiveness were identified and are summarised at the end of this document.

5. Methods

5.1. Evidence was found using a protocol-driven systematic review methodology (The Cochrane Statistical Methods Group, 2011)□. This identified 30 relevant papers on topics including health psychology, substance use, common mental health disorders, and violence. Quality of data was overall moderate. Data were analysed following a realist synthesis (Pawson, Greenhalgh, Harvey, & Walshe, 2004)□.

6. Effectiveness of PCDSSs

6.1. PCDSS impact on mental health outcomes was mixed and unrelated to target disorder. Many studies found no impact, which could be explained by a lack of effect on clinician behaviour: generally they did not use it.

6.2. Where PCDSSs were used by clinicians, treatment tended to be cheaper overall, more effective at reducing symptoms, introduced sooner, and led to faster clinical responses than treatment as usual. Use of these tools rarely resulted in poorer outcomes (see **8. Risks**).

6.3. Overall the results suggested algorithms were effective at improving outcomes when they were used, particularly through improved detection of symptoms and prescription of more effective treatment.

7. Feasibility in practice

7.1. Feasibility was defined as adherence to PCDSS assessment and treatment recommendations in 75% of documented cases. Adherence was reported in nine studies. Feasibility was clearly demonstrated in one study and was negative or unlikely in seven studies.

7.2. These results suggest PCDSSs as generally deployed are not feasible in practice due to lack of voluntary use.

8. Risks

8.1. Risks from the use of PCDSSs

- 8.1.1. **Patient-treatment compatibility:** Algorithms are only as good as the data put into them, and research is limited on what works for whom. PCDSSs are therefore less able to match treatment to individual client preferences than they are diagnoses. For instance, Jenssen et al.'s (2016) PCDSS recommended 165 people for a guideline-based tobacco-cessation program, yet none attended.
- 8.1.2. **Competing interests:** As the purpose of a PCDSS is to alter clinical practice, it will at the very least challenge existing procedures. As many services operate under constrained resources, a tool is likely to compete for those resources with other interests. A tool can therefore lead to 'inappropriate' decisions for a given organisation, such as recommending treatment with limited availability or diverting resources away from Key Performance targets.
- 8.1.3. **PCDSS' understanding of risk factors is limited:** There is less evidence on risk factors or contraindications for treatment, so these are often missing in a PCDSS model. This increases the risk of a given individual receiving inappropriate treatment, particularly when risk is linked to local context. For instance regional laws may differ on whether abuse can be disclosed to a third party without consent, meaning a tool that 'did not know where it was' could potentially recommend illegal action.
- 8.1.4. **Identifying a need and specifying action for resolution creates a moral obligation to address that need:** Friction occurs when client needs are outside the immediate remit of a service. Take a PCDSS that assesses the likelihood of risk for depression and certain cancers. A mental health worker uses this tool and identifies a high risk for cervical cancer in an otherwise healthy client. The clinician has no professional mandate to recommend investigations for cervical cancer, no training to counsel the client, and may not be resourced to make a referral, yet there is a clear risk of preventable harm. This causes ethical and professional dilemmas, which are more likely in multi-disciplinary PCDSSs.
- 8.1.5. **"Availability of good tools alone does not ensure good craftsmanship or clinical judgment"** (Nagpaul, 2008, p. 60): PCDSSs rely on clinicians' skills in gathering information and applying the tool appropriately. For instance, the Suspected Abuse Tool lists "unexplained decreases in bank account" as a marker of financial abuse. However, it is up to the skill of the clinician to illicit enough data to decide whether a given instance is indeed 'unexplained' to the extent it represents abuse. Two practitioners could thus come to different conclusions under similar circumstances.
- 8.1.6. **Cultural insensitivity:** Cultural rules are rarely explicitly built into a PCDSS, and bias is more likely in diverse populations, whether based on age, ethnicity, gender, diagnosis, etc. This can lead to recommendations at odds with particular cultures, and applies equally to clients, practitioners, and organisations. Hunter et al. (2016) cite an example of a PCDSS failing to be adopted because it did not, or was seen to be unable to, take into account the underlying values of the host organisation.
- 8.1.7. **Difficulty accounting for complexity:** PCDSSs are less able to process exceptional, complex, or co-morbid cases as there is less research data to build valid decisional models. Recommendations can be potentially harmful if certain criteria are not accounted for, such as referring someone to a physiotherapist to manage chronic pain when the client also has significant psycho-social issues. First, the tool may not consider who could be best to manage mental health issues, and second, the physiotherapist may be 'over-burdened' by a complex referral.

8.2. Risks to PCDSS uptake

- 8.2.1. **Recommendation is not precise enough:** Human-based decision-making involves more than a simple binary output, even when the judgment itself is between 'yes' or 'no', e.g. how certain is it this recommendation is correct? If 'no', then what? Trust in recommendations is more difficult without this transparency, so practitioners are more likely to reject the tool.
- 8.2.2. **System compatibility with existing routines and hardware:** This applies particularly to time, as many practitioners feel constricted by existing time pressures in their practice. Where PCDSSs are seen as taking time away from important activities, uptake is likely to be low.
- 8.2.3. **Clinicians trust themselves to make decisions more than they trust PCDSSs:** Clinicians tend to believe they make better decisions than algorithms, making them less likely to use a PCDSS or comply with recommendations. Benbenishty found two thirds of clinicians "showed an overwhelming reliance on themselves as decision makers and were almost insulted at the prospect of consulting a computer for support" (pg. 201). Yet the tool in question agreed with professionals more than they agreed with each other.
- 8.2.4. **Insufficient time for training:** Clinicians who were unfamiliar with the PCDSS were less likely to use it. This also worked as a function over time, where occasional use was associated with reducing competence and overall poorer adoption.
- 8.2.5. **Justifying the PCDSS to service users:** Clients may be unwilling to engage with a PCDSS, such as completing online questionnaires, without a clear rationale. Clinicians may be asked to provide such justification, further increasing time pressures in consultations
- 8.2.6. **Ergonomic issues with use, e.g. navigation, speed, and intuitiveness:** Tools that are not 'user-friendly' are harder to use, and therefore more likely to be rejected. Design issues can impact both service users and clinicians if the former needs to interact with the system; Buckingham et al. (2015) noted some clients needed practitioners to assist them with the PCDSS, meaning their usefulness could partly depend on the availability of clinicians. This would offset potential gains from the tool.
- 8.2.7. **What is accepted as important evidence differs:** Tools can be rejected for not including sources of information considered relevant by clinicians, whether or not this is backed by research.
- 8.2.8. **'Considered useful' is not the same as 'will be used':** Many studies received feedback from clinicians that overall the system was thought to be valuable; however, very few guidelines were followed during testing. E.g. Olfson (2003) discovered the majority of clinicians thought the tool helped them diagnose depression, even though their rate of detection was unchanged. These instances strongly argue perceptions of usefulness should not be relied upon as an indicator of uptake or absolute utility in future research.
- 8.2.9. **Implementation lacks (senior) support:** Lack of managerial commitment to implementation usually means a tool will have "died immediately" after the conclusion of a study, regardless of inherent utility (Benbenishty & Treistman, 1998, p. 202).
- 8.2.10. **Making decisions about how decisions are made:** Many studies found decisions were made very differently in practice versus research. If this disagreement between tool and clinician is not settled, the PCDSS is more likely to be seen as lacking appropriate utility and rejected. However, it is unlikely to be resolved in every instance. Research by Barnett

(2002) and Hunter (2016) found even priorities between stakeholders conflicted, including within the same organisation.

9. Conditions for effectiveness

9.1. Involvement of stakeholders early in the PCDSS design process is important to improve the chance of PCDSS use:

- 9.1.1. Uptake of PCDSSs into clinical practice was often hampered by factors beyond its ability to make sound evidence-based decisions. PCDSSs need to be successful on a number of fronts in order to improve chances of being used routinely, including:
1. incorporating the values, priorities, and existing procedures of the organisation in which it is deployed into its model of decision-making. Consider referral criteria for relevant local services, waiting times for treatment, and giving precedence to Key Performance areas.
 2. integrating evidence sources valued by stakeholders—which is usually the expertise of the stakeholders themselves. As a minimum this should include managers and clinicians (to enhance uptake), ideally includes service users (to improve acceptability of recommendations), and could also extend to commissioners and specialist interest groups. This is in addition to research evidence.
 3. being ‘user-friendly’. Clinicians (and clients) are rarely motivated to invest significant time or energy to use an optional system, and the less cognitive effort required to complete a task the better the uptake. Time is a particular barrier to adoption: the longer it takes to learn a system the higher the chance it is dropped beforehand.
 4. feeling ‘owned’ by stakeholders. This enhances a sense of trust in the tool and its decisions, and that the PCDSS is more ‘like me’. It is important for key stakeholders to feel invested in the tool in order to improve adherence, prioritise resources for its use, and continue development as needed.
- 9.1.2. Each area is better addressed in the earliest stage of the tool design process and periodically revisited after development. Studies where these factors were only looked at after the initial research-based model had been established had more difficulty integrating them. Involving stakeholders directly in the design is the easiest way to target these four factors. This can be done using established practices from user-centred design, such as stakeholder workshops, usability testing, and good communication processes (Sharon, 2012).

9.2. PCDSSs improve outcomes for services, clients, and clinicians in the following ways:

- 9.2.1. PCDSSs are most effective for services when they able efficiently use limited resources. If a tool makes a recommendation beneficial to clients, but that depletes service resources overall, ultimately the tool is untenable. However, when care can be matched to particular client circumstances and problems, it is more likely for better care to be provided for the same investment as treatment as usual. This argues PCDSSs should prioritise conditional models of treatment (‘under circumstances x, treat with y, else z’), such as stepped care and personalised treatment approaches. Where possible, contraindications for treatment should also be incorporated.
- 9.2.2. PCDSSs can also benefit services when they highlight additional gaps in decisional need, as long as these can be addressed as efficiently. Suitably incorporating other teams’ requirements into the PCDSS model will also decrease friction at points of contact between

services, for example by increasing the number of appropriate—and therefore accepted—referrals.

9.2.3. Service users profit most from PCDSSs when they are more quickly allocated the care that most effectively addresses their problem of concern. Faster identification occurs when tools can identify symptoms through client feedback, e.g. via a questionnaire, that otherwise the practitioner would not be aware of (for instance asking about symptoms of ADHD is not routine in most primary care centres, so is generally identified only when problems advance and become obvious). However, identification will only be effective if the client accepts the referral. Motivation for treatment is enhanced when the PCDSS is matched to a context that makes sense to the client, such as a paediatrician suggesting they stop smoking because it could harm their child.

9.2.4. Finally, clinicians themselves can benefit from using a PCDSS when it causes them to reflect on their decision-making. As a minimum, this occurs when the clinician sees a tool's recommendation, suggesting PCDSSs should always produce a visible decision. Ideally the decisional model of the PCDSS should be transparent, as this will improve clinician understanding of their own decisional process, as well as enhance their ability to explain it to clients.

9.3. Impact on mental health outcomes depends on clinician behaviour, organisational support, and evidential integration:

9.3.1. Most PCDSSs are constructed on the basis they impact mental health outcomes primarily through making more evidence-based decisions than a clinician. This is incorrect. Most PCDSSs that fail to improve outcomes—despite being evidence-backed—do so in the first instance because clinicians ignore it. This can be for a variety of factors including poor usability or low trust. Behaviour change can be facilitated when clinicians view PCDSSs as valuable tools they work with (rather than being imposed).

9.3.2. The second biggest facilitator of impact is support from the organisation in which the PCDSS is used, usually by senior management. They are important in prioritising resources for the (continued) use of the new system and sustaining motivation to develop the tool over time as new research, procedures, and needs arise.

9.3.3. This support is useful in contributing to the evidence base informing the decisional model. Evidence must come from a range of sources, including managerial priorities. Several studies found considerable resistance to their tools when local, organisational, and professional evidence was not built into their tools. This not only increases the perception of poor utility, but reduces effectiveness by being poorly adapted to its environment. Models based only on research evidence are therefore less likely to have an impact on outcomes.

9.4. It is more important to make valued decisions than decisions that are right according to the research base:

9.4.1. Making the 'right' decision in an absolute sense requires adopting a single point of view to the exclusion of all others, and in mental health care is thus inherently conflictual, given the number of stakeholders involved. For instance treating a person's anxiety disorder may be right for that client, but may not be right for the service when it diverts resources from someone more vulnerable. Thus the different 'rights' need to be weighed up and compromises made, otherwise the tool will most likely be ineffective and rejected.

9.4.2. Decisional models should therefore focus on maximising value to the various key stakeholders (client, clinician, and organisation), which includes but is not exclusively based

on research. What is valuable can be determined through suitable stakeholder involvement. Flow of resources is a common area of value, and is readily addressed by most PCDSSs.

9.5. PCDSSs are more likely to be used and adhered to when they are trusted as decision makers. Trust is related to risk, information, transparency, discretion, and personalisation:

9.5.1. Decreasing perception of risk can come from reallocating responsibility for decisions away from the clinician (people are more likely to take risks if they are not held responsible for poor outcomes) and by communicating good outcomes.

9.5.2. PCDSSs are seen as more trustworthy when they incorporate valued sources of evidence. Including stakeholder expertise can enhance perceptions of trust, as the tool by its nature becomes more 'like me'. This evidence and the way it is used in decisions should be as transparent as possible, especially to clinicians, as this reduces uncertainty (which is anathema to trust). However, if the model output is too complex, the clinician will generally not try and understand it and transparency will be effectively lost.

9.5.3. Feeling in control is important to enhance trust and reduce perception of risk. Practitioners are more likely use a PCDSS if they can adapt its recommendations, although this means more deviations from guidelines. Having decisions imposed on them by a tool, especially without transparency, decreases perceptions of control, and clinicians will naturally try and restore the balance by ignoring the PCDSS. Such discretion also helps personalise decisions to client and context, where PCDSSs are generally at a disadvantage. PCDSSs should demonstrate personalisation where they are able, and allow practitioners to weigh in where they are not.

9.6. PCDSSs are most effective at improving mental health outcomes when they function as part of a mutually symbiotic relationship:

9.6.1. PCDSSs work at their best when able to draw on expertise from multiple sources. All this expertise does not have to be built directly into the algorithm; indeed this would make the model more complicated, harder to understand, and appear less trustworthy as a result. It is often enough to provide opportunities in the design of the tool for input from clinicians, clients, etc. so they can correct any errors the tool makes. PCDSSs are more likely to make mistakes in unusual or complex cases, where human decisions are more valuable. Humans on the other hand make more mistakes when predicting outcomes in general. Integration of expert, contextual, and research knowledge should still be incorporated into the PCDSS where possible, but not at the expense of transparency, comprehension, or opportunities for discretion.

9.7. PCDSSs are more likely to improve mental health outcomes when they are matched to specific contexts and problems:

9.7.1. When PCDSSs are used in a general service to assess for specific problems, they will more often identify people for who treatment is inappropriate. For instance, few people attending a primary care clinic will have symptoms of and desire treatment for Obsessive Compulsive Disorder, while more people attending a mental health centre will. PCDSS usefulness can thus be increased when it is matched to the setting.

9.7.2. PCDSSs can be more effective when they target specific problems, such as feeling low or fatigue, rather than diagnoses, such as depression. This wastes fewer resources treating problems that are not present (e.g. not everyone with post-traumatic stress has nightmares), and can lead to improvement more quickly by targeting symptoms of concern.

10. References

- Barnett, S. R., dosReis, S., & Riddle, M. a. (2002). Improving the management of acute aggression in state residential and inpatient psychiatric facilities for youths. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41(8), 897–905. <http://doi.org/10.1097/00004583-200208000-00007>
- Benbenishty, R., & Treistman, R. (1998). The development and evaluation of a hybrid decision support system for clinical decision making: The case of discharge from the military. *Social Work Research*, 22(4), 195–204.
- Buckingham, C. D., Adams, A., Vail, L., Kumar, A., Ahmed, A., Whelan, A., & Karasouli, E. (2015). Integrating service user and practitioner expertise within a web-based system for collaborative mental-health risk and safety management. *Patient Education and Counseling*, 98(10), 1189–1196. <http://doi.org/10.1016/j.pec.2015.08.018>
- Caulfield, K. J. (2012). *Relationships among Self-Efficacy, Health Beliefs and the Self-Management Behaviors of Healthy Eating and Physical Activity among Adults with Type 2 Diabetes. Relationships Among Self-efficacy, Health Beliefs & the Self-management Behaviors of Healthy Eating & Physical Activity Among Adults With Type 2 Diabetes.*
- Hunter, D. J., Marks, L., Brown, J., Scalabrini, S., Salway, S., Vale, L., ... Payne, N. (2016). The potential value of priority-setting methods in public health investment decisions: qualitative findings from three English local authorities. *Critical Public Health*, 1596(July), 1–10. <http://doi.org/10.1080/09581596.2016.1164299>
- Jenssen, B. P., Shelov, E. D., Bonafide, C. P., Bernstein, S. L., Fiks, A. G., & Bryant-Stephens, T. (2016). Clinical Decision Support Tool for Parental Tobacco Treatment in Hospitalized Children. *Applied Clinical Informatics*, 7(2), 399–411. <http://doi.org/10.4338/ACI-2015-12-RA-0169>
- Lie, D. A., Lee-Rey, E., Gomez, A., Bereknyei, S., & Braddock, C. H. (2010). Does cultural competency training of health professionals improve patient outcomes? A systematic review and proposed algorithm for future research. *Journal of General Internal Medicine*. <http://doi.org/10.1007/s11606-010-1529-0>
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* <http://doi.org/10.1037/11281-000>
- Nagpaul, K. (2001). Application of Elder Abuse Screening Tools and Referral Protocol: Techniques and Clinical Considerations. *Journal of Elder Abuse & Neglect*, 13(2), 59–78.
- Olfson, M., Tobin, J. N., Cassells, A., & Weissman, M. (2003). Improving the detection of drug abuse, alcohol abuse, and depression in community health centers. *Journal of Health Care for the Poor and Underserved*, 14(3), 386–402. <http://doi.org/10.1177/1049208903255454>

- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). Realist synthesis - an introduction. *ESRC Research Methods Programme*, (January 2004), 1–46. Retrieved from <http://discovery.ucl.ac.uk/180102/>
- Ruiz, J. M., Hamann, H. A., Garcia, J., & Lee, S. J. C. (2015). The psychology of health: Physical health and the role of culture and behavior in Mexican Americans. In *Mexican American children and families: Multidisciplinary perspectives. BT - Mexican American children and families: Multidisciplinary perspectives*.
- Sharifi, M., Adams, W. G., Winickoff, J. P., Guo, J., Reid, M., & Boynton-Jarrett, R. (2014). Enhancing the electronic health record to increase counseling and quit-line referral for parents who smoke. *Academic Pediatrics*, *14*(5), 478–484. <http://doi.org/10.1016/j.acap.2014.03.017>
- Sharon, T. (2012). *It's Our Research. It's Our Research*. <http://doi.org/10.1016/C2010-0-66450-9>
- The Cochrane Statistical Methods Group. (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0. Cochrane Handbook for Systematic Reviews of Interventions*. <http://doi.org/Available> from www.cochrane-handbook.org.
- Tversky, A., & Kahneman, D. (1973). *Judgment under uncertainty: Heuristics and biases. Judgment under uncertainty: Heuristics and biases*.
- Wilkinson, S., & Himstedt, K. (2008). Establishing an innovative model of nutrition and dietetic care for a mental health service through collaboration with non-nutrition healthcare workers. *Nutrition and Dietetics*, *65*(4), 279–283. <http://doi.org/10.1111/j.1747-0080.2008.00310.x>