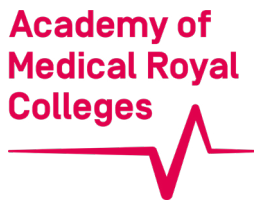May 2022

Academy of Medical Royal Colleges

A paper on the principles and attributes of ensuring the credibility of health information in social media

Preface

All of us, patients, the public more widely and indeed clinicians, increasingly seek information on health-related issues via social media platforms. Indeed, for many, social media platforms are their only source of timely health information.

Currently there are no consistent processes for assuring the quality of the vast quantity of health information available on-line. While much health information is helpful and of high quality, there is plenty which is not. As an organisation representing health professionals the Academy of Medical Royal Colleges has deep concerns about the quality of such information particularly potentially harmful "fake" information. The Academy is clear that providing, independent and objective assurance about the quality of health information and health information providers in a digital age is hugely important and needs to be done well.

We were therefore pleased to be approached by Google to discuss this issue and particularly to be invited to provide an independent commentary on the paper from the United States of America's National Academy of Medicine, which looked at the issue in the context of the USA health landscape.

This is not a formal Academy policy statement. Rather it is an objective commentary and discussion paper which has been supported by the Academy Council and is intended to be the start and not the end of a discussion.

We are keen to continue this discussion with the full range of stakeholders – patient groups, clinicians, Government and social media providers themselves – to seek to collectively develop a system which can provide the public with effective assurance of the quality of health information they may access through social media platforms. There is a pressing need to progress this work which will be in the best interests of everyone seeking reliable, trustworthy answers.

Professor Dame Helen Stokes-Lampard
Chair, Academy of Medical Royal Colleges

Contents

Introduction

In late 2021 Google approached the Academy of Medical Royal Colleges in the United Kingdom [the Academy] to discuss how health information and providers of health information on social media platforms [SMPs] available in the UK could be quality assured to ensure their accuracy. Their particular interest was in information on YouTube which is owned by Google.

Previously Google had approached the National Academy of Medicine [NAM] in the USA to discuss this same issue. In July 2021 NAM published an individually authored discussion paper *Identifying Credible Sources of Health Information in Social Media: Principles and Attributes*. Associated information relating to the project can be found on NAM's website.

The Academy was clear that the task of providing objective assurance about the quality of health information and health information providers in a digital age is hugely important. As an organisation representing health professionals we have deep concerns about the quality of much of the information available to patients and the public through SMPs and particularly potentially harmful "fake" information. This therefore is a task that needs to be done.

We are aware that this is also a concern to NHS England which has also held discussions with Google. NHSE has produced a standard for creating health content which could help assess and assure the process of how health information is produced.

However, exactly how the whole process that might be done in a UK context and who might be involved is a subject for wider discussion.

Google, having brought the NAM paper to the attention of the Academy, felt it would be helpful if the Academy could provide an independent commentary on it, particularly focussing on its applicability [or not] to the context of healthcare in the UK.

This task has been undertaken on a voluntary unpaid basis, completely independently from Google irrespective of whether the Academy or Colleges have any role in future assurance processes. As with the original NAM paper, this is not a formal policy position from UK medical royal colleges but rather a commentary and discussion paper which has been supported by the Academy Council, which comprises the presidents of the medical royal colleges, in February 2022.

The National Academy of Medicine discussion paper

Identifying credible sources of health information in social media: principles and attributes

NAM stated that the purpose of their project was "*to identify preliminary definitions of "authoritative" sources of health information and the criteria by which these sources derive and maintain their authority.*"

The paper says credible sources of health information in social media should be science-based, objective, transparent, and accountable. The paper also outlines attributes of credible sources of health information, such as providing citations for information shared and evidence to justify claims, acknowledging limitations and evolution of knowledge, and prioritizing accessibility and equitable access to information.

The paper then goes on to discuss additional actions beyond identification of source credibility which could be taken by Social Media Platforms.

NAMs methodology

The process for the production of NAM's paper was

— Advisory Group: An expert advisory group appointed by the NAM gathered information and deliberate in order to author a discussion paper for publication in *NAM Perspectives*

— Public Comment Opportunity: Initial definitions, principles, and questions were posted on NAM's webpage for public comment

— Information-Gathering Webinar: The NAM hosted a public webinar on April 5, 2021, heard feedback from experts, stakeholders, and the public

— Feedback Synthesis: A summary of feedback received during the public comment period and webinar was posted on their webpage

— Final Discussion Paper: An individually-authored, peer-reviewed paper containing guidance for consideration by social media platforms was published in Summer 2021.

It is important to recognise that NAM was clear that the paper represents the views of the individual authors following the project, rather than formal NAM policy. The report is not an affirmation (or otherwise) of the actions of any individual social media platform.

Key NAM conclusions

The NAM paper has two main strands

— Defining a methodology for determining source credibility

— Associated actions that are required to assure the quality of health information on social media platforms.

This paper looks at these two strands, with particular focus on the first one, offers a commentary on the findings and the new NHSE Standard for creating health content and then suggests some general conclusions which should underpin any further work and some questions which need to be addressed.

Determining source credibility — scope

At the outset NAM was clear that it was considering only the credibility of sources of health information not content. Therefore, judgements would be on the credibility of the providers of health information, not on the accuracy or quality of any specific individual content. The rationale for this was essentially pragmatic in that assuring the quality of all content would be an unachievably vast task. However, assuring the credibility of source providers will inevitably be a proxy measure and while credible providers should produce credible information, as NAM says, "*The credibility of a source is, at most, an indicator of information quality and by no means a guarantee*". The NAM paper discusses ways this may happen with both official institutions and academic journals.

The authors further limit their consideration in the paper to government and non-profit organisations (including non-profit news sources that share health information) and does not consider individuals (e.g. independent clinicians with Facebook pages) or for-profit companies.

The rationale for these exclusions is that individuals require separate analysis because they lack the organisational infrastructure that is the basis of the authors' approach to source evaluation. For-profits have a unique set of financial interests that, likewise, require a separate assessment.

Academy commentary

NAM's limitation of its scope is pragmatic and very understandable. But it does raise problems — as NAM itself is the first to acknowledge.

However, the decision to seek to assure the credibility of sources rather than content seems entirely sensible. Attempting to assure the quality of all information content on social media platforms would be an unachievable, Sisyphean task.

However, source credibility is only a proxy and there remains the possibility that information of insufficient quality is produced by otherwise credible organisations, whether they be government, professional or academic organisations. The second issue is ensuring currency of information from credible sources as the science progresses with information becoming outdated.

Assuring the credibility of sources is not the only action that can be taken to ensure quality of information. Potential actions by SMPs or others, including action to taken on demonstrably false information, are considered later.

But, in summary, it would certainly seem sensible in the UK to swiftly adopt a similar process as the US, by focusing on assuring sources/providers of information.

The context of profit and non-profit is obviously different in the UK at least in terms of health providers. There are, for profit organisations who provide health information in the UK (e.g. pharmaceutical or device manufacturers]. As a matter of principle the aim should be to ensure that all sources, irrespective of whether they are governmental, not-for profit or for profit are considered as to their credibility as sources.

The Academy's Patient/Lay Committee Chair made the point, endorsed by the Academy Council, that the NAM paper can be seen as favouring expert professional opinion at the expense of information from patient groups and third sector organisations. Patient information and the evidence behind it is not solely scientific based but can include experiential understanding and social science perspectives. This is not to say that such information should not be produced to robust standards but the evaluation criteria need to ensure that this approach is incorporated and not excluded.

The issue of individuals is different. There is plenty of excellent health information posted on social media platforms by individual clinicians, scientists or others, but there is much which is not. The NAM argument that individuals are highly unlikely to have the organisational infrastructure to meet any evaluation criteria is strong.

It may be that in taking this forward, consideration could be given to some form of self-assessment criteria for individuals, but at present it is hard to see how a robust, valid process to assess individual credibility could be devised.

Determining source credibility — principles and attributes

The NAM paper then goes on to develop "foundational principles" to guide identification of credible sources of health information in social media. These are, in essence, the criteria by which judgements on source credibility could be made:

— **Principle 1: Science-Based**
Sources should provide information that is consistent with the best scientific evidence available at the time and meet standards for the creation, review, and presentation of scientific content.

— **Principle 2: Objective**
Sources should take steps to reduce the influence of financial and other forms of conflict of interest or bias that might compromise or be perceived to compromise the quality of the information they provide. This principle acknowledges that all sources have COIs or inherent biases. However, in order to be considered credible, sources should strive to separate the presentation of health information from profit motives and other biases (e.g. political).

— **Principle 3: Transparent and Accountable**
Sources should disclose the limitations of the information they provide, as well as conflicts of interest, content errors, or procedural missteps. The final principle acknowledges the fallibility of both organizations — which cannot eliminate COI and errors — and science itself.

The paper supplements these principles with a series of "attributes" for each principle which can be used to judge whether the core principles are being met. The attributes are attached as Annex A.

Academy commentary

Both the foundational principles and the attributes appear sound. Any set of criteria for judging source credibility has to enable a clear-cut assessment to be made as to whether a provider can demonstrate it produces reliable evidence-based high quality health material according to a transparent set of standards for objective, evidence-based, quality assured information.

The principles and attributes seem to meet that requirement well. While there could be some tweaks or amendments for use in a UK context, the Academy believes that they could be used in conjunction with the NHSE Standard for creating health content (discussed later) in the UK.

Determining source credibility – the process

The NAM paper then considers how the process of determining source credibility based on the principles above might work in practice.

Importantly, the paper identifies several categories of sources as being subject to pre-existing, standardised vetting mechanisms or accreditation that indicate general alignment with the foundational principles and attributes. These sources could include health providers, government organisations and academic journals. Therefore, sources in these categories could be considered at the outset as likely to be credible.

However, these mechanisms are not likely to be an exact match for the principles and attributes and reflect varying degrees of rigour. It should not be assumed that everyone will automatically accept the credibility of government, academic or provider sources.

NAM states that accreditation/vetting mechanisms do not evaluate an organisation's credibility as a source of health information in social media. Rather, they affirm the general credibility of an organisation in its role as a provider of a specific service, such as education or health care. Such organisations may not always adhere to the authors' credibility principles and criteria.

Therefore, accreditation/prior vetting is an imperfect proxy for the evaluation of an organisation's credibility as a source of health information and should be viewed as a preliminary indicator. The question is, therefore, whether these formalised vetting systems can be taken as a sufficiently strong proxy to afford an assumption of credibility or whether further assessment is required and what that should be.

This of course also leaves all those organisations which are not subject pre-existing, standardised vetting mechanisms that align with the authors' principles and attributes. The paper provides a useful table [Page 31] of types of non-profit organisations [in the US context] not subject to standardised vetting mechanisms with observation of their credibility likelihood. The types of organisations include professional bodies and societies, think tanks, Foundations, patient or disease advocacy groups and health industry groups.

NAM state that these groups would need be assessed on a standardised set of data. The issue is whether those tasks should be undertaken by social media providers themselves or independently. NAM states that "*a particularly promising possibility for consideration by SMPs would be to outsource QA functions to an independent third party, either pre-existing or created for this purpose*".

Academy commentary

The status of governmental bodies, regulation and accreditation is very different in a UK context to that of the US. However, the principles the paper raises would seem to be the same in the UK context.

Organisations pre-judged as credible sources

The question is whether there is a range of bodies whose status and/or robust and transparent regulatory regime merits an automatic or preliminary assumption of source credibility.

It would seem sensible that this is the case. For practical and logical purposes it seems right that a range of UK organisations such as Government Arm's Length Bodies, NHS provider organisations, academic journals, universities should be recognised as credible providers of health information.

Healthcare organisations in England, for example, are all registered and regulated by the CQC. Similarly to the accreditation process in the US, CQC is assessing the fitness of organisations as healthcare providers. It is not assessing the credibility of their processes for producing health information, although it is something that CQC could consider as part of its overall assessment.

This raises of the question of what happens when an organisation is judged inadequate in its service provision. Also, CQC regulates health and care organisations from large hospitals to single handed primary care medical and dental practices and there is a question of whether the same level of automaticity applies across all health providers.

Government departments, national NHS bodies, charities, and academic institutions and journals are not regulated in the same way as healthcare providers, but do have clear and transparent governance and accountability mechanisms. Consideration would need to be given as what requirements or criteria would be expected of these bodies.

As in the US, partial or automatic acceptance of credibility of particular types of organisations will not be infallible and the question remains as to whether any further form of verification either on a universal, selective or random basis is required. Accompanying self-assessment relating to the NHSE standard for creating health content may provide a solution.

Ensuring credibility of all health information on social media platforms may never be possible. What can be done will, of necessity, be done incrementally rather than in a single move. It would be sensible and pragmatic to start with those organisations where an initial presumption of credibility is reasonable.

It is therefore suggested that further work is undertaken at the earliest opportunity to identify the criteria by which organisations in the UK could be included in that category of automatic credibility and the additional requirements, if any, of those organisations. This needs to be a collaborative initiative between representatives of social media platforms and relevant stakeholders (Government, healthcare providers, health regulators, academic and professional organisations and the public).

Organisations not pre-judged as credible

As in the US, this then leaves the question of how organisations who would not be considered to be within the category of automatically credible, through dint of their regulatory/governance arrangement, should be assessed. It is very important not to make any implicit assumption that organisations not included in the regulated/statutory governance are somehow not likely to be credible sources. This group will include many organisations producing information to the highest standards.

Again, we believe that the NAMs proposal for some form of assessment against standardised data sets provided by the organisation is sensible. How this could be done and who would do it remains for detailed consideration although we believe using the NHSE Standard for creating health content (see later) would be sensible and provide a consistency of approach. The Academy recognises NAM's point about the value of social media platforms outsourcing this task to an independent organisation.

It also seems that at this stage, and probably in perpetuity, that this credibility assessment which effectively 'kite marks' the provider has to be on a voluntary participation basis. Organisations which want to be publicly recognised as having been judged to be a credible source of information would apply and provide the required information. Forcing all providers of information to go through the process before they could post information on SMPs would be both unmanageable and very much against the ethos of open access which is at the core of social medial platforms. However, we believe that it is important that SMPs clearly signal to the public where providers have been assessed as credible and where they have not.

Consideration needs to be given to the specific question of news/media outlets. Many will be credible sources, but some may not adhere to desirable standards — and all can make errors. Whether the same criteria should be applied to media outlets as other organisations, whether there should be specific criteria or whether existing mechanisms for assuring accuracy should be used needs to be considered.

Finally, there remains the issue of individuals as providers of health information. Essentially, a similar process of voluntary self-assessment/registration could be developed. However the data set required would have to be different as individuals, as has been noted, could not have the same organisational infrastructure as organisations. There is a question of whether any process for individual providers of health information

could actually be robust enough to be of any real added value. Again, any compulsory requirement for individuals to undertake the process is likely to be unmanageable whether or not it is considered desirable. We believe the issue of individuals as providers of health information should be considered separately at a later phase.

The Academy believes that what is appropriate and practical for both non pre-judged organisations and individuals has to be the subject of further details work and discussion by social media platforms and key stakeholders.

Parallel strategies to supplement source assessment

The remainder of the NAM paper looks at other strategies required in parallel to source credibility assessment.

These include fall into three categories:

1. Quality of information

Management of Misinformation

The paper states that SMPs should maintain parallel strategies to address such false and inaccurate information, as well as sources that deliberately promulgate such information.

Health Literacy, Culturally Competent Communication, and Community Relationships

The paper argues that regardless of the eventual system for elevating credible sources and high-quality information, consumers will still make their own judgments about which sources and information to trust. This will be based on a range of issues relating to their personal and social circumstances.

"Cultural competence" refers to the ability to interact effectively with diverse audiences by recognizing and responding to variations in social, cultural, and linguistic needs.

The paper says there should be major efforts in public education and education of content providers in how to ensure information address these issues.

Source Self-Regulation and Consumer Evaluation

Rather than waiting for evaluation by SMPs, sources of health information that wish to be considered credible should take proactive steps to apply science-based, objective, and transparent and accountable principles to their institutional practices and presentation of information. Consumer evaluation can play a part in this but is not sufficient in itself.

2. Ethical and public health considerations

Control of Information

This considers concerns at the perceptions of potential censorship or limiting the autonomy of information providers.

Structural Bias

Ensuring equality, diversity and inclusion.

Financial Conflict

Ensuring that any form of accreditation/kite-marking does not create its own financial conflicts although it is recognised that financial gain and enhanced influence may be unavoidable collateral effects of designation as a credible source in social media channels.

Feasibility and Appropriateness of SMPs' Role

The paper states that the complexity of SMPs' interests merits careful consideration of their role as moderators of health and other crucial public information. Although the authors believe that SMPs should take a proactive role across several dimensions, as outlined in this paper, government regulation and delegation to independent third parties should also be considered as potential supplementary approaches.

3. Public health considerations

Health Equity

A system to elevate credible sources — and thereby increase access to high-quality health information — must be designed to support health equity, as well as information equity, and not cement existing inequities. SMPs should also consider digital literacy and strategies to address equity in access to high-quality digital information — a challenge referred to as the "digital divide".

Contribution to Public Health Research

SMPs can be important partners in improving public health, but only if they agree to share data (e.g. backend data, algorithms and use engagement metrics, content moderation processes) with researchers.

Academy commentary

The Academy does not disagree in principle with the comments made by the authors and indeed would strongly support the call around management of misinformation, health equity, the role of SMPs and public health research. We believe all are worthy of further discussion between SMPs and wider stakeholders.

However, they are not immediately applicable to the issue of assuring the credibility of information sources and therefore we have not addressed them in detail.

While not specifically referred to in the NAM paper, in terms of content assurance there is a model whereby there are a number of trusted and credible 'super-users' of SMP which act as the first point of verification for other users. If one or a number of such users endorse or support specific information, it can be recognised as a form of recommendation or kite mark. Such a group could also act collectively to review materials. Of course, there are questions as to who is designated as a "super-user" and on what basis. However, it is certainly a concept to explore further.

WHO Roundtable: global principles for identifying credible sources of health information on social media

In December 2021 the World Health Organization (WHO) held an [online consultation meeting](#) to discuss global principles for identifying credible sources of health information on social media and applicability of the NAM principles globally.

The meeting concluded that while the "science-based" principle was generally more straightforward to apply in a global context, the principles of being "objective" and "transparent and accountable" posed more challenges. There was agreement among participants that these principles could mean different things in different settings and that factors such as culture, language, the influence of the source of information and political views within the country can affect what is seen as credible.

Participants flagged that disclosing financial and non-financial interests could be problematic and interpreted in different ways. A suggested attribute to include in the principles of "transparent and accountable" was "does not sell a product"

In conclusion, the participants noted that there would be some challenges in applying the principles and attributes in different contexts, there was consensus that they could be applied globally.

The Roundtable also made suggestions for social media providers themselves around issues such as media literacy training and content moderation.

Academy commentary

The WHO's broad support for the NAM principles is welcome. We fully recognise the points made about the challenges of interpretation and application in different global contexts. While there are certainly differences in context between the US and the UK, there are also many areas where there is an alignment of cultural context. Application of the NAM principles is therefore probably more straightforward in the UK than it might be in some countries.

While the WHO report was considered only when much of the Academy's work was complete there is nothing in it which would lead the Academy to amend the conclusions of its commentary.

NHS standard for creating health content

As stated at the outset NHSE (previously NHSX) has produced a standard for creating health content. The standard outlines best practice in creating good quality health content and are designed for any organisation that produces health and care information.

Adherence to this standard could be used as a template to assess and assure the credibility of organisations in producing health information. NHSE would certainly expect all NHS organisations to follow the standard in producing health content.
The Academy believes that an approach which combines that set out in this commentary on NAMs proposals and the NHS standard would provide a solid basis for beginning the task of assuring the credibility of health information providers on social medial platforms.

Conclusions

The Academy believes,

— There is no doubt that the task of providing objective assurance about the quality of health information and health information providers in a digital age is hugely important.

— We are under no illusions that this is a complex task which is bound to be, at least in its initial stages, incomplete.

— For practical reasons, any process should initially be restricted to seeking to assure the credibility of providers of health information as a proxy for content itself.

— For practical reasons, any process should be restricted to seeking to assure the credibility of organisational providers of health information, not individuals.

— It is reasonable to define a range of statutory or regulated organisations whose status could confer an automatic or at least preliminary assumption of source credibility.

— Work is needed between social media providers, Government and other stakeholders to determine the criteria for organisations included in that first group.

— Determination is needed as to what further assessment, if any, is required for these organisations. This could incorporate self-assessment against the NHSE standard for creating health content.

— Work is needed to determine the process by which organisations not included in the statutory/regulated cohort would be assessed as credible sources. This could involve assessment against a transparent data set, again aligned to the NHSE standard for creating health content. This will show:

  — The nature of the provider and any of its potential conflicts of interest

  — How the provider researches/collates/gathers or otherwise generates its information and source material in terms of process, evidence requirements, quality assurance and independence of opinion

  — How the provider ensures regular review of its information to ensure it is current.

— Specific consideration needs to be given to the position of news/media outlets.

— The process of assurance should ideally be run independently of social media providers themselves.

— Work is needed between social media providers, Government and other stakeholders to determine precisely how that process for assuring source credibility would work for all organisations.

— This should include exploration of the concept of "super-users".

— Work is needed between social media providers, Government and other stakeholders on a series issues not directly related to source credibility but around the quality of content on topics such as management of misinformation, health equity, the remit of SMPs and public health research.

Annex A

Foundational principles and attributes of credible sources of health information

Foundational Principle	Attributes
Science-Based: Sources should provide information that is consistent with the best scientific evidence available at the time and meet standards for the creation, review, and presentation of scientific content.	— Acknowledges the limitations and evolution of knowledge (e.g., early or incomplete knowledge, as seen in the COVID-19 pandemic; small sample size; correlation versus causation, etc.) — Clearly labels information with the date it was last updated and strives to reassess and update content — Demonstrates subject-specific expertise (i.e., consistent and well-regarded contributions in a given field) — Links to and is linked to by other credible sources [a] — Provides citations for information shared and evidence to justify claims — Synthesizes information from multiple sources, rather than a single source — Uses a consensus process to develop the information shared [b] — Uses peer review or another form of content review to vet information before sharing [c]
Objective: Sources should take steps to reduce the influence of financial and other forms of conflict of interest (COI) or bias that might compromise or be perceived to compromise the quality of the information they provide.	— Keeps health information separate from financial, political, or ideological messages — Maintains independence from funders [d] — Separates lobbying activities from health information (or does not engage in lobbying) — Does not include advertisements with relevant health information (or does not host advertisements at all) [e]
Transparent and Accountable: Sources should disclose the limitations of the information they provide, as well as conflicts of interest, content errors, or procedural missteps.	— Discloses financial and nonfinancial conflicts — Discloses relevant policy positions and lobbying activities — Follows FACA regulations or similar transparency policies [f] — Posts public corrections or retractions — Prioritizes accessibility and equitable access to information — Provides a mechanism for public feedback — Shares data, methods, or draft recommendations

a. For example, an organization could seek public comments on an interim set of health guidelines before finalizing and sharing the information more broadly.

b. A consensus process involves assembling a group of experts with diverse perspectives who assess a body of evidence and deliberate in order to arrive at an opinion or guidance that reflects the consensus of the group.

c. A peer review process involves sharing the draft of a publication or other product with reviewers who have expertise or experience in the given topic and can provide feedback as to the product's accuracy, balance, and appropriateness.

d. For example, an academic journal could maintain editorial independence (i.e. sole authority over published content) from the organization that funds it.

e. For example, an organization might host an advertisement for a cancer drug but keep this advertisement separate from the information it shares about cancer.

f. FACA stands for the Federal Advisory Committee Act, which established requirements for committees that advise the federal government. These requirements include public access to meetings and meeting notes, as well as summaries of expenditures (https://www.gsa.gov/policy-regulations/policy/federal-advisory-committee-management/advice-and-guidance/the-federal-advisory-committee-act-faca-brochure)

Source: Raynard S. Kington, Stacey Arnesen, Wen-Ying Sylvia Chou, Susan J. Curry, David Lazer, and Antonia M. Villarruel. *Identifying Credible Sources of Health Information in Social Media: Principles and Attributes. Discussion Paper*. July 16, 2021. Available at https://nam.edu/identifying-credible-sources-of-health-information-in-social-media-principles-and-attributes/