Articles

Pilot testing of a new electronic diagnostic support tool (AiD-DST) designed to identify the cause(s) of delirium

Eamonn Eeles[1], Peter Worthy[1], Andrew Teodorczuk[1], Paven Kaur[2], Nadeeka Dissanayaka[1]

[1] University of Queensland, [2] Queensland Health

Delirium Communications

Introduction

Delirium is frequently multifactorial, and causes are often missed in clinical practice. The Aetiology in Delirium - Diagnostic Support Tool (AiD-DST) was developed to improve recognition of the causes. We undertook an evaluation of an electronic version of AiD-DST.

Methods

A development and evaluation life cycle of improvement was used. In phase 1, alpha testing among the development group evaluated technical performance of AiD-DST. In phase 2, we performed a cycle of beta testing among junior doctors to assess impressions of AiD-DST using Think Aloud methodology. We grouped responses into themes and made changes to AiD-DST by the development group accordingly. In phase 3, usability and acceptance of AiD-DST was assessed using the mHealth App Usability Questionnaire (MAUQ).

Results

In phase 1, software issues were identified, and modifications made. In phase 2, feedback was obtained from 29 junior doctors. Three cycles of feedback were obtained. The number of items identified after each cycle were 20, 12 and 7, respectively. Content was grouped into themes of; 'style and grammar', 'formatting', 'IT', 'missed diagnosis' and 'other concerns.' In phase 3, 20 participants completed MAUQ questionnaire. Overall, the average score was 6.36 (SD=0.8) with 7 as the highest attainable score. This translates to agreement up to strong agreement concerning usability of AiD-DST.

Conclusion

After a process of optimisation, AiD-DST has been shown to be a usable and potentially useful diagnostic support tool to help junior doctors identify cause(s) of delirium. An implementation study is planned.

> "*By becoming interested in the cause, we are less likely to dislike the effect.*"
> ⁻ Dale Carnegie, How to Win Friends and Influence People

INTRODUCTION

Delirium is a common disorder in older and frail hospitalised older adults and is associated with many adverse health outcomes.[1] The causes of delirium outlined in the multifactorial causation model are protean and multiple.[2,3] For instance, in a study of 105 patients who screened positive for delirium in the emergency department, the median number of aetiology categories identified was two. Nearly one-third of delirious patients in this study identified 3 or more causes for delirium.[4] In a review of delirium causation, polypharmacy, psychoactive medication use, and physical restraints were leading factors in medical patients.[5] Illness severity, alcohol misuse, catheterisation and iatrogenesis are other associations.[5] However, heterogeneity in aetiological representation across studies is common. Variability in delirium causes arises from a consistent pool of factors whose individual pre-eminence is dictated by contextual factors. These include psychomotor subtype, clinical setting, and patient age.[5] For instance, causes between younger and older patients are different, with lung and cardiac disease more common in older patients in a liaison psychiatry setting.[6] In the oldest-old, these aetiologies shift again. In a sample of 3,076 patients aged >80 years, 42% of whom developed delirium, renal failure, intracranial haemorrhage and pleural effusions were putative causes.[7] Even when the same patient conditions are studied, there is little replication in causality.[8]

These associations also fail to encompass the diversity of potential causalities in any given patient where the multifactorial nature of the condition is the only reliable rule. Time-pressured, junior physicians often work out of hours and make decisions about delirium when there is the least available senior support.[9] Clinicians frequently resort to diagnostic heuristics, which lead to acceptance at face value of singular and more convenient diagnoses, such as urinary tract infection, as a cause for delirium.[10] Inability to follow instructions on the part of the patient owing to drowsiness or extremes of psychomotor activity might pose additional barriers to clinical evaluation.[11] Such barriers may lead to oversimplification of diagnosis and marginalisation of important aetiological factors, with clinical ramifications.[9] For instance, a greater than 24 hours delay to treatment of delirium causes in an ICU setting led to a quadrupling of mortality.[12] This high mortality remained even after adjustment for other covariates.

In response to this common and serious challenge, we designed an algorithm that would capture the representative and multifactorial cause(s) of delirium that could be used to identify clinical causes within a medical setting. The Aetiology in Delirium - Diagnostic Support Tool (AiD-DST) algorithm was developed from first principles by a multi-professional team with a track record of innovation and cognitive tool development in delirium[13–15] (see appendix 1). The algorithm intended to follow the clinical flow of assessment by the physician, capture common causes with the greatest yield and use Bayesian principles of probability to improve efficiency.[16] For instance, if the patient were immobile, then the algorithm would direct towards screening of pressure ulcers which would otherwise be omitted in a mobile patient.[17] The algorithm was then refined based upon open feedback from senior geriatricians into a minimum of eight diagnostic steps (see appendix 1).

AiD-DST was tested in the clinical setting in patients with a diagnosis of delirium. Sensitivity of the algorithm is 88.8% with a specificity of 71.8% against research-grade diagnosis of the cause(s) of delirium.[13,18] AiD-DST content has been transposed into a progressive web application. The final electronic version was built on Microsoft Blazor technology and is available for use on any device (see appendix 2). Our objective was to refine the AiD-DST in light of feedback and evaluate usability among front-line junior medical staff.

METHODS

This was a multi-stage cycle of refinement of AiD-DST between February and December 2020. The study was in keeping with principles of an agile development and evaluation lifecycle.[18] This involves user experience design, development, alpha and beta testing. In phase 1, we performed technical testing of AiD-DST within the development group. In phase 2, junior doctors evaluated AiD-DST for content and acceptability. In phase 3, junior doctors assessed the usability and usefulness of AiD-DST according to a standardised survey.
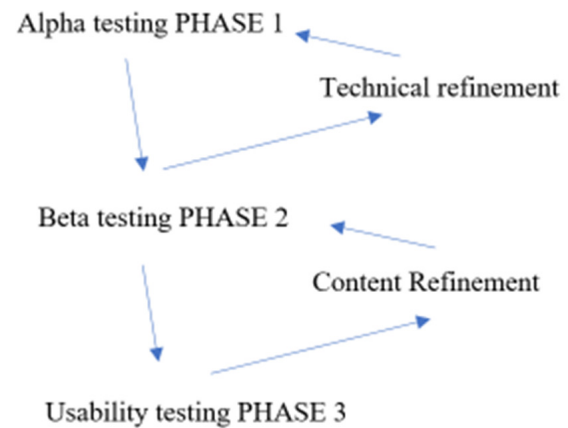

Figure 1. Study pathway

Informed assent was obtained from study participants before taking part in the study. The Local Research and Ethics committee approved the study, Project ID 56923.

PARTICIPANTS

The junior doctors were from a medicine department of a single Metropolitan Hospital site. The medicine department comprised acute medicine and subacute services.

EXPOSURE (PHASE 1, 2 AND 3)

Phase 1: We performed alpha testing of AiD-DST by the development group (comprising 3 geriatricians and a software engineer) to establish technical operating characteristics of AiD-DST. Consensus-approved changes were implemented into AiD-DST (Figure 1).

Phase 2: Junior doctors performed beta testing. AiD-DST was introduced at face-to-face core medical teaching with an invitation to any interested juniors to participate in the study. PW, an advanced trainee in geriatrics, followed up with a phone call or face-to-face invitation. PW provided a brief training session on the technical aspects of how to access and use the AiD-DST. Participants were asked to 'think-aloud' while accessing and using the AiD-DST. Think Aloud methodology is a valid way of evaluating processes, particularly at the human-design interface.[19] This involved the participant expressing impressions to the investigator while running the AiD-DST task. Participants were asked to conceive of a prior patient with a delirium diagnosis to help with context. We intended the total session to last just a few minutes (less than ten). PW transcribed comments from participants during completion of the Think Aloud task. A sample size of between 5 and 20 participants were chosen for each cycle.[20] Affinity mapping was used to group feedback into representative categories.[21] We adopted a consensus-building approach within the development group to identify and sanction modifications (Figure 1). A cycle of review was repeated using the same method until there was no further modifiable feedback.[22] Modifications were incorporated into AiD-DST by our software engineering team af-

Table 1. Themes created by affinity mapping and corresponding changes made to AiD-DST by the expert group.

Affinity mapping Themes	Changes by reference group
Style and Grammar	
"Not to use medical abbreviations e.g. SOL, hyperCa" "Provide an introduction to the app"	Most abbreviations have been removed: "SOL" is now "space occupying lesion" 'hyperCa' is now "hypercalcaemia" An introductory page is in development.
Formatting	
"All the questions were leading in favour of abnormal findings and then it was confusing to have the questions reversed. E.g., 'is y normal?... tripped you up."	Style has been changed so that all are leading questions in preference of abnormal findings e.g., "is the gait abnormal?" as opposed to "is the gait normal?"
Information Technology	
"Would be nice to modify an answer rather than restart"	Currently AiD-DST remains unchanged such that if a mistake is made the user needs to restart. There are technical barriers to modification of this.
Clinical signs insufficiently sensitive	
"Misses infective symptoms/signs if the answer is no to fever but older patients may not mount a fever and have an infection'"	Phrasing changed to be broader to capture infection "or infective symptoms/signs and include immunosuppression?"
Excluded Diagnoses	
"B12 deficiency" "Constipation"	Vitamin B12 deficiency was added. Constipation was added as a potential cause/ contributor to all cases.
Other concerns	
"Important to contextualise and remember it doesn't replace clinical judgement"	Consideration of a disclaimer added to AiD-DST "intended as an aid to and not a replacement for clinical judgement"

ter each cycle. To avoid a change in psychometric properties of AiD-DST, the overall structure of the algorithm remained unchanged, and removal of any diagnoses reported in the original study was avoided. Junior doctors who were, at the time of the study, within the same medical team as an investigator were excluded.

Phase 3: A survey of junior doctors was conducted over four weeks using the mHealth App Usability Questionnaire (MAUQ) (see appendix 3).[23] The MAUQ is a validated usability questionnaire designed to evaluate health applications. The questionnaire has 18 statements divided into subscales: ease of use, interface, satisfaction, and usefulness. A seven-point ordinal satisfaction scale is used to grade responses with higher values representing agreement and lower values. Doctors were invited to complete a paper-based version of MAUQ, and anonymous responses were collected. A pragmatic sample size of 20 respondents was intended based upon recommendations from the literature concerning usability studies.[24]

ANALYSIS

Usability data was reported using descriptive methods, including mean and standard deviation.

RESULTS

In phase 1, the development group identified episodic freezing of the algorithm. The software was therefore changed from a powerapp format to Microsoft Blazor technology. The technical issues had resolved upon reassessment.

Phase 2 comprised three rounds of feedback that sampled 29 out of a total of 63 eligible junior doctors within an internal Medicine. Eighteen respondents (62%) were in either years 1 or 2 of postgraduate medical training and 11 were year 3 up to specialist registrar equivalent. Three cycles of feedback were obtained with changes made to AiD-DST at each stage. The number of items identified after each cycle were 20, 12 and 7, respectively. Feedback was subsequently categorised into themes that included 'style and grammar', 'formatting', 'information technology', 'insufficiently sensitive signs', 'excluded diagnoses' and 'other' (see Table 1). An example of an issue identified in 'grammar' was confusion over the use of medical abbreviations. The development group recommended that nomenclature be provided in full without acronyms. The software engineer implemented these changes. Under 'possible missed diagnoses', we subsequently added vitamin B12 deficiency and constipation/ urine retention to the list of putative delirium causes. No diagnostic causes were removed from AiD-DST.

In phase 3, twenty junior doctors (13 men, 7 women), used the refined version of AiD-DST. Number of years into postgraduate medical training was a mean of 4.1; SD=1.9. Response rate was 100% for 16 of the 18 questions. With 7 as the highest attainable score, average overall score was 6.4 (SD=0.8), which represents 'agreement' to 'strong agreement' concerning usability. Impressions of usability persisted across the subscales with at least 'agreement' in ease of use, interface and satisfaction and usefulness with average scores of 6.5 (SD=1.1), 6.5 (SD=0.9) and 6.1 (SD=1.1), respectively. Absence of an internet connection and how the app handled mistakes were individual items

Table 2. Showing usability performance of the AiD-DST according to MAUQ.

	Statements	Mean score (SD)
	EASE OF USE	
1	The app was easy to use.	6.7(0.7)
2	It was easy for me to learn to use the app.	6.7(0.8)
3	The navigation was consistent when moving between screens.	6.7(0.8)
4	The interface of the app allowed me to use all the functions (such as entering information, responding to reminders, viewing information) offered by the app.	6.5(1.0)
5	Whenever I made a mistake using the app, I could recover easily and quickly.	5.1(1.8)
	Overall ease of use score	6.5(1.1)
	INTERFACE AND SATISFACTION	
6	I like the interface of the app.	6.3(1.0)
7	The information in the app was well organised, so I could easily find the information I needed.	6.5(0.9)
8	The app adequately acknowledged and provided information to let me know the progress of my action.	6.2(1.0)
9	I feel comfortable using this app in social settings.	6.8(0.5)
10	The amount of time involved in using this app has been fitting for me. *	6.9(0.3)
11	I would use this app again	6.5(0.8)
12	Overall, I am satisfied with this app.	6.4(0.8)
	Overall Interface and satisfaction score	6.5(0.9)
	USEFULNESS	
13	The app would be useful for my healthcare practice.	6.3(0.9)
14	The app improved my access to delivering healthcare services.	6(1.0)
15	The app helped me manage my patients' health effectively.	6.1(1.0)
16	This app has all the functions and capabilities I expected it to have.	6.3(1.0)
17	I could use the app even when the Internet connection was poor or not available.	5.6(1.6)
18	This mHealth app provides an acceptable way to deliver healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment.	6(1.0)
	Overall Usefulness score	6.1(1.1)

In this questionnaire, 1 - strongly disagree, 2 – disagree, 3 – somewhat disagree, 4 – neither agree nor disagree, 5 – somewhat agree, 6 – agree, 7 – strongly agree

*The average time spent on AiD-DST according to the server database was 40 seconds (range: 1-169 seconds).

within the MAUQ with a score of less than 6 (somewhat agree). These were also only answered by 25% and 50% of respondents, respectively. This recurred as an area of critique in open commentary (see table 1). No adverse events or technological failure was reported.

DISCUSSION

In this study we showed the process of refinement and optimisation of an electronic diagnostic support tool, AiD-DST, and report on its high usability rating among junior doctors. Recognising the causes of delirium can be a challenge, even in experienced hands. Junior doctors need guidance through the multifactorial web of causality in delirium. The AiD-DST is the first and only digital diagnostic support tool that may help provide a sophisticated analysis for the benefit of junior front-line clinicians who may not always have a consultant physician to hand. This might be particularly the case out of hours when delirium is most prevalent. The perception of usefulness and acceptance of AiD-DST by ju-

nior doctors is a promising foundation for clinical implementation.

Algorithms have been used to diagnose other neurological conditions such as stroke syndromes with a favourable positive predictive value of up to 0.91 (95%CI:0.71-1.0).[25] Smartphone-based tests are moderately accurate in differentiating mental health disorders in adulthood.[26] Moreover, smartphone-based tests have been successfully validated in screening for delirium and discriminating it from dementia.[27]

One of the challenges for an algorithm for delirium aetiology was the inter-patient variability in causality, creating a potential tension between inclusivity and ergonomics. Our original study identified key points of difference in patient presentation that would help focus clinical enquiry while avoiding redundancy in questions. For instance, presence of a fever would prompt a deeper dive into possible sources that would be omitted beyond screening questions in an afebrile patient. Once AiD-DST was validated as a diagnostic tool, it was therefore imperative to identify that it lived up to the intention of also being usable and helpful in

clinical practice once in its desired format of a smartphone device. A process of further refinement prior to evaluation identified several modifiable factors and corresponding improvements were made by the expert reference group. This finalised version of AiD -DST was shown to be usable and acceptable by a sufficiently sized sample of junior doctors drawn from the acute medical setting.[23]

Limitations of this study include the possibility of response bias. Firstly, the investigator was an advanced trainee in geriatrics within internal medicine at the time. While not negated altogether, this risk of influencing junior doctors was minimised to a degree by excluding junior doctors from the researcher's own medical team. Further studies will attempt to capture the user profiles in more depth and explore utility in experienced providers.

Second, demographic information concerning the respondents was limited. It is possible that certain groups were not represented. However, the junior medical workforce is diverse and there was a high level of consistency in the responses, as indicated by the small variance.

Third, notes were taken by only one researcher and not transcribed independently, or audio recorded. We acknowledge this potential threat to reliability of feedback that informed changes to AiD-DST. Due to the intensity of work-load for junior doctors and competing clinical priorities, we decided that bringing the research to the doctor would be the most feasible methodology. Despite the necessarily lean research method, prima facie feedback appeared constructive and much of it was actionable by the reference group. There was minimal substantive change to the content from the original version. Now that AiD-DST has shown promise for use in clinical practice, we anticipate exploring the potential to adapt AiD-DST for virtual evaluation of patients in cases of COVID-19, where delirium is both common and contact precautions create impediments to clinical assessment.[28]

The AiD-DST is a first-of-its-kind electronic support tool that requires minimal training, is usable and is deemed helpful by target users for clinical practice. AiD-DST can be accessible when junior doctors most need it in the format they are most likely to use. With a demographic shift in front-line clinician work practices, there is an imperative to embrace digital technologies that can improve acumen, and delirium aetiology is rightly an emergent focus. Further validation and implementation studies are planned.

REFERENCES

1. Wilson JE, Mart MF, Cunningham C, et al. Delirium. *Nat Rev Dis Primers*. 2020;6(1). doi:10.1038/s41572-020-00223-4

2. Inouye SK, Charpentier PA. Precipitating factors for delirium in hospitalised elderly persons. Predictive model and interrelationship with baseline vulnerability. *JAMA*. 1996;275(11):852-857. doi:10.1001/jama.1996.03530350034031

3. Laurila JV, Laakkonen ML, Tilvis RS, Pitkala KH. Predisposing and precipitating factors for delirium in a frail geriatric population. *Psychosom Res*. 2008;65(3):249-254. doi:10.1016/j.jpsychores.2008.05.026

4. Cirbus J, MacLullich AMJ, Noel C, Ely EW, Chandrasekhar R, Han JH. Delirium etiology subtypes and their effect on six-month function and cognition in older emergency department patients. *Int Psychogeriatr*. 2019;31(2):267-276. doi:10.1016/j.jpsychores.2008.05.026

5. Hshieh TT, Inouye SK, Oh ES. Delirium in the Elderly. *Psychiatr Clin North Am*. 2018;41(1):1-17. doi:10.1016/j.psc.2017.10.001

6. Grover S, Ghormode D, Ghosh A, et al. Risk factors for delirium and inpatient mortality with delirium. *J Postgrad Med*. 2013;59(4):263-270. doi:10.4103/0022-3859.123147

7. Marquetand J, Bode L, Fuchs S, Ernst J, von Känel R, Boettger S. Predisposing and Precipitating Factors for Delirium in the Very Old (≥80 Years): A Prospective Cohort Study of 3,076 Patients. *Gerontology*. 2021;67(5):599-607. doi:10.1159/000514298

8. Magny E, Le Petitcorps H, Pociumban M, et al. Predisposing and precipitating factors for delirium in community-dwelling older adults admitted to hospital with this condition: A prospective case series. *PLoS ONE*. 2018;13(2):e0193034. doi:10.1371/journal.pone.0193034

9. Eeles EM, England R, Armstrong A, Pinsker D, Pandy S, Teodorczuk A. Understanding our patients better will lead to better recognition of delirium: An opinion piece. *Australas J Ageing*. 2018;37(4):241-242. doi:10.1111/ajag.12585

10. Dasgupta M, Brymer C, Elsayed S. Treatment of asymptomatic UTI in older delirious medical in-patients: A prospective cohort study. *Arch Gerontol Geriatr*. 2017;72:127-134. doi:10.1016/j.archger.2017.05.010

11. FitzGerald JM. Delirium clinical motor subtypes: a narrative review of the literature and insights from neurobiology. *Aging Ment Health*. 2018;22(4):431-443. doi:10.1080/13607863.2017.1310802

12. Heymann A, Radtke F, Schiemann A, et al. Delayed treatment of delirium increases mortality rate in intensive care unit patients. *J Int Med Res*. 2010;38(5):1584-1595. doi:10.1177/147323001003800503

13. Eeles E, Huang L, Dakin L, et al. Development and validation of an aetiology in delirium diagnostic support tool. *Age Ageing*. 2020;34(6):547-551. doi:10.1093/ageing/afaa269

14. Eeles E, Gunn H, Sutt AL, et al. e-Screening revolution: A novel approach to developing a delirium screening tool in the intensive care unit. *Australas J Ageing*. 2018;37(2):147-150. doi:10.1111/ajag.12525

15. Lin HS, Eeles E, Pandy S, Pinsker D, Brasch C, Yerkovich S. Screening in delirium: A pilot study of two screening tools, the Simple Query for Easy Evaluation of Consciousness and Simple Question in Delirium. *Australas J Ageing*. 2015;34(4):259-264. doi:10.1111/ajag.12216

16. Bayes T. An Essay Toward Solving a Problem in the Doctrine of Chances. *Philosophical Transactions*. 1764;53:370-418.

17. Alderden J, Rondinelli J, Pepper G, Cummins M, Whitney J. Risk factors for pressure injuries among critical care patients: A systematic review. *Int J Nurs Stud*. 2017;71:97-114. doi:10.1016/j.ijnurstu.2017.03.012

18. Wilson K, Bell C, Wilson L, Witteman H. Agile research to complement agile development: a proposal for an mHealth research lifecycle. *NPJ Digit Med*. 2018;1(1):46. doi:10.1038/s41746-018-0053-1

19. Ericsson AK, Simon HA. How to study thinking in everyday life: contrasting think aloud protocols with descriptions and explanations of thinking. *Mind Cult*. 1998;5(3):178-186. doi:10.1207/s15327884mca0503_3

20. Lundgrén-Laine H, Salanterä S. Think-Aloud Technique and Protocol Analysis in Clinical Decision-Making Research. *Qual Health Res*. 2010;20(4):565-575. doi:10.1177/1049732309354278

21. George ML, Maxey J, Rowlands DT, Price M. *Working with Ideas. The Lean Six Sigma Pocket Toolbook*. McGraw – Hill; 2005:27-32.

22. Jake-Schoffman DE, Silfee VJ, Waring ME, et al. Methods for Evaluating the Content, Usability, and Efficacy of Commercial Mobile Health Apps. *JMIR Mhealth Uhealth*. 2017;5(12):e190. doi:10.2196/mhealth.8758

23. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. *JMIR Mhealth Uhealth*. 2019;7(4):e11500. doi:10.2196/11500

24. Faulkner L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*. 2003;35(3):379-383. doi:10.3758/bf03195514

25. Tanigawa M, Kohama M, Nonaka T, et al. Validity of identification algorithms combining diagnostic codes with other measures for acute ischemic stroke in MID-NET®. *Pharmacoepidemiol Drug Saf*. 2022;31(5):524-533. doi:10.1002/pds.5423

26. Jungmann SM, Klan T, Kuhn S, Jungmann F. Accuracy of a Chatbot (ADA) in the diagnosis of mental disorders: comparative case study with lay and expert users. *JMIR Form Res*. 2019;3(4):e13863. doi:10.2196/13863

27. Tieges Z, Stott DJ, Shaw R, et al. A smartphone-based test for the assessment of attention deficits in delirium: A case-control diagnostic test accuracy study in older hospitalised patients. *PLoS ONE*. 2020;15(1):e0227471. doi:10.1371/journal.pone.0227471

28. Rogers JP, Chesney E, Oliver D, et al. Psychiatric and neuropsychiatric presentations associated with severe coronavirus infections: a systematic review and meta-analysis with comparison to the COVID-19 pandemic. *Lancet Psychiatry*. 2020;7(7):611-627. doi:10.1016/s2215-0366(20)30203-0

SUPPLEMENTARY MATERIALS

Appendix 1

Download: https://deliriumcommunicationsjournal.com/article/37365-pilot-testing-of-a-new-electronic-diagnostic-support-tool-aid-dst-designed-to-identify-the-cause-s-of-delirium/attachment/96310.pdf

Appendix 2

Download: https://deliriumcommunicationsjournal.com/article/37365-pilot-testing-of-a-new-electronic-diagnostic-support-tool-aid-dst-designed-to-identify-the-cause-s-of-delirium/attachment/96311.docx

Appendix 3

Download: https://deliriumcommunicationsjournal.com/article/37365-pilot-testing-of-a-new-electronic-diagnostic-support-tool-aid-dst-designed-to-identify-the-cause-s-of-delirium/attachment/96312.doc