Commentary: The best-laid plans: the problems and pitfalls of assessing mild cognitive impairment

Wendy Johnson

Centre for Cognitive Ageing and Cognitive Epidemiology and Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK. E-mail: wendy.johnson@ed.ac.uk

Now is the winter of our discontent...
—Shakespeare, Richard III

The best-laid schemes o’ mice an’ men gang aft agley, An’ lea’e us nought but grief an’ pain, for promis’d joy!
—Robert Burns, To a Mouse

What a frightening thing is the human, a mass of gauges and dials and registers, and we can only read a few and those perhaps not accurately.
—John Steinbeck, The Winter of Our Discontent

The concept of mild cognitive impairment (MCI) was developed with the best of intentions. The term refers to a set of clinically recognizable characteristics that indicate poorer cognitive function than would be expected based on the age and education of the patient, but which are not severe enough to warrant a diagnosis of dementia. The concept is useful, at least in theory, because dementia is inherently a degenerative syndrome. MCI begins mildly but gets steadily, or even increasingly, more severe, and we have few if any effective treatments for most cases. Catching its emergence early or identifying those most vulnerable to MCI would help both in identifying its causes and courses and thus treatments, and in managing cases even before we have effective treatments. The concept of MCI was developed to do just those things.
But MCI is one messy concept! Cognitive function varies considerably in groups of people at any age. It also has many different aspects. Within the population at any given age, these different aspects correlate: people who function well in one area tend to do so in others too, and vice versa. But this is only a tendency. People can and do often function relatively well in one area and not so well in another, for example spatial reasoning and general knowledge, or memory and verbal fluency. To make things worse, there are normative developmental patterns: overall function increases in youth, relatively stabilizes sometime in early adulthood and begins an inexorable decline some time after that. But the various aspects of function reach their normative peaks and begin their normative declines at very different ages and decline at different rates. Moreover, there are individual differences in these rates, but they tend to be only fractions of the individual differences in level of function in any group of people at any one time. This means that even if cognitive tests were completely reliable, there would be less statistical power to detect changes in function than differences in level of function, and population models inevitably must acknowledge the presence of both. And of course cognitive tests are never completely reliable. Their accuracy suffers from state effects such as fatigue, caffeine consumption and test anxiety, not to mention familiarity with the testing materials. This means that even when substantial decline in function is detected in an individual, recovery often seems apparent the next time around.

All of this feeds into diagnosing what we call dementia. The cognitive areas affected range from memory to attention to language to reasoning, and just one or any combination can be affected. The causes vary considerably as well, and range from head injury in relative youth to vascular disease that impedes the flow of oxygen to the brain to the accumulation of plaques and tangles in the brain associated with Alzheimer’s disease to substance abuse to downstream consequences of other neurological diseases such as Parkinson’s. Because dementia is inherently degenerative, cognitive function normally shows such large individual differences at any age, these differences vary with level of education and cognitive function normally declines with age, diagnosis is usually based on level of performance expected normatively for someone of the individual in question’s age and level of education, often supplemented by subjective reports of family members that the individual is struggling to cope with everyday affairs. But of course there are considerable differences in cognitive function even among people of the same age and level of education, and one of the bottom-line criteria for diagnosis is inability to manage day-to-day activities. People with lower levels of peak ability and/or less education tend to reach this threshold much sooner than others because less decline from peak levels is required.

All the complications associated with normal cognitive function and dementia diagnosis feed into defining and diagnosing MCI, as Kremen et al. clearly are aware. They took on two additional challenges. First, as they note, average age in MCI studies runs in the 70s, but MCI can set in well before that. Early identification of this common precursor of dementia would be even better than just detecting MCI, so they estimated its prevalence rates in a community sample of men aged in their 50s. And second, they attempted to estimate the extent of genetic influences on MCI. Their sample had completed a rather extensive battery of cognitive tests, so the various aspects of cognitive function were well represented. Their sample had another advantage as well: the participants had completed a measure of general cognitive function when they were about age 20 years. This should be a good measure of peak general cognitive function.

To meet the challenges they had taken on Kremen et al. made use of five quite different definitions of MCI, all taken from the literature, and broke them out into three subtypes of MCI. To gauge prevalence rates, they regressed age 20-years general cognitive test scores from each of the more specific assessments made in mid life. This approach explicitly recognized that MCI is inherently about decline in function. However, because the general cognitive measure does not address the nuances of the various aspects of cognition, interpretation of these residualized change scores must rely on the tendency for individuals who do relatively well in one area of cognition to do relatively well in all the others, and vice versa. This is only a tendency. Moreover, as Rogosa, Brandt and Zimowski noted long ago, these residualized change scores are subject to many statistical problems and biases, and they do not measure change directly. Rather, they measure something like what the individual’s deviations are from the change that would have taken place if everyone had started out equal, which of course they did not.

Given all this, I found the study’s results difficult to interpret. Prevalence rates for MCI ranged from 2.57% to 64.74% across the various definitions! Some of this variance of course reflects the variance in thresholds inherent in the (very arbitrary) definitions, but much of it also, undoubtedly at least to me, reflects the complications I’ve discussed above. Such a range is epidemiologically useless, and any syndrome or clinical warning sign that captures 64% of the population is likely to be too. It’s also unclear how to apply these prevalence rates to the much more common samples or situations where measures of peak cognitive function are not available.

And then there are the estimates of genetic influence. They reflected only too well the imprecision of measurement of MCI, with ranges from 0 to 0.56 and wide confidence intervals that often included zero. We already know that essentially every human trait that’s measured at all precisely is heritable, to the point that
Eric Turkheimer\textsuperscript{8} christened this the First Law of Behaviour Genetics over a decade ago. We also know that more accurate measurement produces higher estimates of genetic influence,\textsuperscript{9} so I would have been amazed not to see the wide range of estimates this study produced. We know the estimate should not be 0 or 1.00, but what it should be in between is anybody’s guess. Based on this study, I would not recommend following Kremen \textit{et al.}\textsuperscript{6} suggestion that we invest scarce research funding on a genome-wide association study of MCI.

The winter of my discontent with the state of MCI affairs is upon me. What can be done about that? The first step is clearly to come to better agreement on what MCI should represent. Should it be the lower end of the distribution of late-life cognitive function, in general, in any one specific area or in some particular number of specific areas? If either of the latter two, which area(s)? If the latter one, how many? And how far down the distribution? Or should it be a threshold of function level rather than a portion of the distribution? What level? Or should it represent change in function, regardless of level? How can this best be most economically/conveniently tapped, given that prior measures of function are not generally available when a patient presents with a concern about decline in cognitive function? And again, what areas of function and what levels of decline? The Kremen \textit{et al.} study’s\textsuperscript{6} strongest message, to me at least, is that research surrounding MCI will not make much progress until these questions are addressed.

\textbf{Conflict of interest:} None declared.

\textbf{References}

\begin{enumerate}
\item Deary IJ, Johnson W. Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. \textit{Int J Epidemiol} 2010;\textbf{39}:1362–69.
\item Johnson W, Penke L, Spinath FM. Heritability in the era of molecular genetics: some thoughts for understanding what it is and what it is not. \textit{Eur J Personality} 2011;\textbf{25}:254–66.
\end{enumerate}