Self-driving vehicles (AVs) are an emerging technology with the potential to save millions from serious injury and death, as well as enriching our lives in other ways. However, realising this potential will take a long time. This assertion is based on two self-evident truths. One, it will take time to perfect the technology. Two, it will take time to proliferate the technology.

As an avenue to reduce global road fatalities, not just those in (primarily) wealthy and industrialized nations, AVs are a technology of little consequence. Limited or non-existent driver education, unsafe driving cultures and poor road/traffic design are all much more bigger contributors. Improving the predictability of the roads might even be necessary before a self-driving car can even be operated safely and usefully in such an environment.

For the wealthy nations with well-developed roads and traffic systems, AV development is one of the most promising areas of technological development on the horizon. Automating vehicles comes with numerous other benefits for both commercial and personal applications. Increased mobility of the disabled, reducing congestion and making commuting easier and more pleasurable are all very likely outcomes of AV adoption. The promises of improved safety are as much a spearhead for other improvements as they are a justification.

However, the road to the self-driving future is paved with hazards. Whilst tempered now, startup culture in Silicon Valley was permeated with a cavalier attitude towards safety. The Deep Learning models that power some AV systems might be incredibly capable, but we are still not even close to the stage where we understand their decision-making as we do with bespoke and human-coded Finite State Machines. Simulations might go some of the way to alleviating this, but the kinds of situations that might cause undesirable behaviour are the ones that are least likely to be modelled in simulations to begin with.

Lack of regulation could see irresponsible manufacturers shipping products that simply aren’t ready, and lack of appropriate legal framework could either put off or bite users, especially early adopters.

So called “trolley problems” might represent a fractional number of situations that lead to death or serious injury, but coming to a consensus on them as a society may be necessary for wider public acceptance of AVs. The transition phase itself represents a significant risk to human safety. We can’t predict with certainty that AVs will handle novel situations safely. Past experience with humans should lead us to assume someone, somewhere, will lose their life doing something stupid. No amount of well-focused awareness campaigns will reduce integration risk to zero, calibrating humans to AVs and AVs to the wider world will simply take time.

Excessive caution is also a potential risk, albeit one that can only be understood with the benefit of hindsight. A high-profile failure or two could turn the public against AVs, delaying safety benefits that could prevent further harm.

Other more novel risks include the use of an AV as a weapon. If a networked AV was compromised by someone with ill intent, it could be used to harm with a high level of discrimination. Such an attack could be carried out across borders, with anonymity, and appear accidental.

Despite risks outlined I believe that an aggressive adoption strategy in countries capable of supporting such an endeavour are in everyone’s interests. This does not represent a belief that self-driving vehicles will be anything close to perfect, merely that they will be significantly better than humans in a short window of time. Proceed caution with caution.